ENTWICKLUNG UND VALIDIERUNG DES AGGRESSIONS-MOTIV-GITTERs.

Ein Beitrag zur Messung des Aggressionsmotivs und der Aggressions-Hemmung bei Kindern.

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Philosophie
der Philosophischen Fakultät
der Universität des Saarlandes

vorgelegt von
Klaus Peter Burkhardt
aus Bexbach

SAARBRÜCKEN 1986

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41552704_0006

ENTWICKLUNG UND VALIDIERUNG DES AGGRESSIONS-MOTIV-GITTERs.

Ein Beitrag zur Messung des Aggressionsmotivs und der Aggressions-Hemmung bei Kindern.

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Philosophie
der Philosophischen Fakultät
der Universität des Saarlandes

vorgelegt von
Klaus Peter Burkhardt
aus Bexbach

SAARBRÜCKEN 1986

Der Dekan: Prof. Dr. D. Fliedner
Berichterstatter: Prof. Dr. H.-J. Kornadt
Prof. Dr. W. H. Tack
Tag der letzten Prüfungsleistung: 6. Mai 1986

BILDUNGSGANG

12. 12. 1948	geboren in Bexbach
1955 - 1959	Besuch der katholischen Volksschule (Pestalozzi-Schule) in Bexbach
1959 - 1960	Besuch des Realgymnasiums in Neunkirchen / Saar
1960 - 1966	Besuch des staatlich anerkannten, altsprachlichen Privatgymnasiums der SOCIETAS VERBI DIVINI in St. Wendel
1966 - 1968	Besuch des staatlichen, altsprachlichen Gymnasiums in St. Wendel und Abitur
1968 - 1971	Besuch der Pädagogischen Hochschule Saarbrücken und 1. Staatsprüfung für das Lehramt an Grund- und Hauptschulen
1972 - 1978	Studium der Erziehungswissenschaft (Hauptfach), Allgemeinen Psychologie (Nebenfach) und Sozialpsychologie (Nebenfach) an der Universität des Saarlandes und Magisterexamen

I N H A L T

Vorwort

I. THEORETISCHER TEIL

1. Einleitung 2

2. Aggression 5
   2.1 Begriff 5
   2.2 Theoretische Konzepte 10
   2.3 Zusammenfassung 19

3. Diagnostik der Aggression 21
   3.1 Perspektiven diagnostischer Aufgabenstellung 22
   3.2 Meßmodelle und testmethodische Konzepte 24
   3.3 Darstellung einiger ausgewählter Erhebungsverfahren 29
   3.4 Zusammenfassung 78

4. Problemstellung 80

5. Grundlagen der Entwicklung des AM-GITTERs 83
   5.1 Motivationstheorie der Aggression 83
   5.2 Gitter-Technik 89

6. Bestandteile des AM-GITTERs 93
   6.1 Situationsseitige Bestandteile 94
   6.2 Personseitige Bestandteile 96
   6.3 Item 97

7. Vorgabe und Funktionsweise des AM-GITTERs 98
   7.1 Vorgabe 98
   7.2 Funktionsweise 102

II. EMPIRISCHER TEIL

1. Voruntersuchungen zur Entwicklung des AM-GITTERs .. 108
1.1 Vorüberlegungen zur Operationalisierung 108
1.1.1 Bildsituationen 108
1.1.2 Aussagen 110
1.2 Erste Schritte zur Konstruktion 114
1.2.1 Experimentelle Prüfung der Bildsituationen 114
1.2.1.1 Fragestellung 114
1.2.1.2 Methode 115
1.2.1.3 Ergebnisse 118
1.2.1.4 Diskussion 120
1.2.2 Entwicklung konstruktthematischer Aussagen 122
1.2.2.1 Wahl der Kategorien 122
1.2.2.2 Expertenrating 124
1.2.3 Entwicklung und Funktion neutraler Aussagen 125
1.3 Zusammenstellung und Prüfung einer Testvorform . 126
1.3.1 Vorbefragung 126
1.3.2 Testzusammenstellung 128
1.3.3 Durchführung der Untersuchung 128
1.3.4 Prüfung der Bildsituationen und Aussagen . 130
1.3.4.1 Fragestellung 130
1.3.4.2 Methode 131
1.3.4.3 Ergebnisse 132
1.3.4.4 Diskussion 133
1.3.5 Dimensionalität der konstruktthematischen Aussagen 134
1.3.5.1 Fragestellung 134
1.3.5.2 Methode 136
1.3.5.3 Ergebnisse 136
1.3.5.4 Diskussion 139
1.3.5.5 Ein alternatives Modell 140

1.3.6 Probeweise Berechnung von Skalen 144
1.3.6.1 Wahl des Modells 144
1.3.6.2 Zusammenstellung der Skalen 145
1.3.6.3 Schätzung einiger Skalenparameter . 147
1.3.6.4 Dimensionalität der Skalen 149
1.4 Zusammenfassung 152

2. Hauptuntersuchung und Konstruktion des AM-GITTERs .. 154
2.1 Einleitende Gesichtspunkte 154
2.2 Zusammenstellung einer neuen Testvorform 155
2.2.1 Selektive Maßnahmen 156
2.2.2 Konstruktive Maßnahmen 157
2.3 Durchführung der Untersuchung 159
2.3.1 Stichprobe 159
2.3.2 Datenaufnahme 161
2.4 Prüfung der Bildsituationen und Aussagen 162
2.4.1 Fragestellung 162
2.4.2 Methode 162
2.4.3 Ergebnisse 162
2.4.4 Diskussion 163
2.5 Dimensionalität der Aussagen 164
2.5.1 Vorüberlegungen 164
2.5.2 Ausgangshypothesen 168
2.5.3 Methode 169
2.5.4 Ergebnisse 169
2.5.5 Diskussion 173
2.5.6 Generalfaktormodell 179
2.6 Selektion der endgültigen Skalen 185
2.6.1 Vorüberlegungen und Wahl des Modells 185
2.6.2 Richtlinien der Zusammenstellung 186
2.6.3 Standard-Skalen 189
2.6.3.1 Fundamentale Skalen 189
2.6.3.2 Abgeleitete Skalen 189
2.6.4 Differentielle Skalen 192

2.7 Reliabilität der Skalen des AM-GITTERs 195
2.8 Dimensionalität der Skalen des AM-GITTERs 197
2.8.1 Ausgangshypothesen 197
2.8.2 Methode 198
2.8.3 Ergebnisse 198
2.8.4 Diskussion 202
2.8.5 Theoretische Festlegung der Dimensionalität der STD-Skalen des AM-GITTERs 204
2.8.5.1 Voraussetzungen 204
2.8.5.2 Vektorielle Einordnung der STD-Skalen in das Generalfaktormodell 206
2.9 Zusammenfassung 212
3. Untersuchungen zur Validität des AM-GITTERs 219
3.1 Einleitende Gesichtspunkte 219
3.2 Meßwertestabilität 221
3.2.1 Vorüberlegungen 221
3.2.2 Ausgangshypothesen 223
3.2.3 Methode 224
3.2.4 Ergebnisse 227
3.2.4.1 Stabilität individueller Kennwerte 227
3.2.4.2 Stabilität der Meßwerteverteilungen 229
3.2.5 Diskussion 231
3.3 Diskriminante und konvergente Validität 237
3.3.1 Vorüberlegungen 237
3.3.2 Vergleichstests 240
3.3.3 Methode 243
3.3.4 Präzisierung der Zusammenhangsannahmen ... 244
3.3.4.1 PFT-Variablen 244
3.3.4.2 PFK-Variablen 247
3.3.4.3 LMG-Variablen 249
3.3.5 Ergebnisse 251
3.3.5.1 Vergleichsmessung PFT 251

3.3.5.2 Vergleichsmessung PFK 252
3.3.5.3 Vergleichsmessung LMG 253
3.3.6 Weiterführende Analysen 255
3.3.6.1 Vorüberlegungen 255
3.3.6.2 Methode 258
3.3.6.3 Ergebnisse 259
3.3.6.3.1 Vergleichsmessung PFT ... 259
3.3.6.3.2 Vergleichsmessung PFK ... 262
3.3.6.3.3 Vergleichsmessung LMG ... 264
3.3.7 Diskussion 266
3.4 Korrelation mit Außenkriterien 273
3.4.1 Vorüberlegungen 273
3.4.2 Ratingmethode 276
3.4.2.1 Entwicklung 276
3.4.2.2 Skalen 279
3.4.2.3 Dimensionalität der Skalen 281
3.4.3 Zusammenhangsannahmen 284
3.4.4 Methode 285
3.4.5 Ergebnisse 287
3.4.6 Weiterführende Analysen 288
3.4.6.1 Vorüberlegungen 288
3.4.6.2 Methode 290
3.4.6.3 Ergebnisse 292
3.4.7 Diskussion 295
3.5 Attribution 301
3.5.1 Vorüberlegungen 301
3.5.2 Fragestellung 303
3.5.3 Methode 304
3.5.4 Ergebnisse 306
3.5.5 Diskussion 308
3.6 Zusammenfassung 312
4. Gesamtzusammenfassung 318
5. Literatur 322
6. Anhang 335

Vorwort

Mit der vorliegenden Arbeit berichten wir über die Entwicklung eines psychodiagnostischen Verfahrens zur Messung des Aggressionsmotivs und der Aggressions-Hemmung bei Kindern. Aufgrund des Erkundungscharakters der Arbeit haben wir den Geltungsbereich des Verfahrens von vornherein relativ eng abgesteckt. Er erstreckt sich auf Jungen im 4. Grundschuljahr (etwa 9 bis 11 Jahre). Sondergruppen, wie etwa Sonderschüler oder in stationärer Therapie befindliche Kinder, sind dabei nicht berücksichtigt.

Die genannte Aufgabenstellung wurde erstmals im Rahmen einer Magister-Arbeit in Angriff genommen (BURKHARDT 1978). Aufgrund der ermutigenden Ergebnisse dieser ersten Untersuchungen (vgl. KORNADT 1982 a, S. 130 f.; KORNADT, ZUMKLEY & BURKHARDT 1982) bestand die berechtigte Hoffnung, bei Fortführung der Bemühungen ein Testverfahren zu erhalten, das im Ergebnis gleichermaßen mit unseren theoretischen Vorgaben zur Natur und Funktionsweise kindlicher Aggressivität als auch mit psychometrischen Richtlinien in Einklang gebracht werden konnte. Dies führte in der Folge dann zur Weiterentwicklung, endgültigen Konstruktion und erstmaligen Prüfung der Gültigkeit des in diesem Bericht vorgestellten Verfahrens.

Inhaltlich haben wir den vorliegenden Bericht in 2 Teile untergliedert: Im theoretischen Teil geben wir zunächst einen skizzenhaften Überblick über die vielgestaltige Verwendungsweise des

Begriffs "Aggression" sowie einige der wichtigsten, zu diesem Gegenstandsbereich entwickelten theoretischen Konzepte. Damit verfolgen wir lediglich die Absicht, dem Leser die Einordnung unseres eigenen theoretischen Ansatzes sowie die des davon abgeleiteten Testverfahrens zu erleichtern. Den gleichen Zweck soll eine nachfolgend vorgestellte, kleine Auswahl von Verfahren zur Erfassung kindlicher Aggressivität erfüllen, die sich zum Teil auf einzelne der o.g. Konzepte als theoretischen Hintergrund beziehen. Im Kern berichten wir sodann über unser eigenes, motivationstheoretisches Konzept zur Aggression sowie über den testmethodischen Ansatz, den wir zur Entwicklung unseres Verfahrens gewählt haben. Wir ergänzen diese Ausführungen um erste Konturen zur Funktionsweise und Induktionslogik unseres Motivtests.

Im empirischen Teil fassen wir zunächst das Wesentliche der bis zum Abschluß der Voruntersuchungen ermittelten Ergebnisse zusammen und halten uns dabei an eine bewußt kurze, überblicksartige Art der Darstellung, soweit das allgemeine Vorgehen als auch Einzelheiten dazu anderenorts (BURKHARDT 1978; KORNADT, ZUMKLEY & BURKHARDT 1982) bereits detailliert vorgestellt wurden. Über bislang noch nicht vorgestellte Analysen zu den in den Voruntersuchungen erhobenen Daten berichten wir allerdings in der gebotenen Ausführlichkeit. Ein Teil der anschließend vorgestellten Hauptuntersuchung war bereits Gegenstand einer fachrichtungsinternen Veröffentlichung (KORNADT, ZUMKLEY & BURKHARDT 1982). Wie bereits bei

den Voruntersuchungen, so stellen wir auch hier die dort referierten Ergebnisse mit dar, um einen geschlossenen Überblick über die insgesamt durchgeführten Arbeiten und deren Ergebnisse zu gewährleisten. Demgegenüber gehen wir in dieser Arbeit insbesondere Fragen der Dimensionalität der Meßwerte unseres Verfahrens sowie Fragen der Meßbarmachung bestimmter Teilaspekte des Aggressionsmotivs wie auch der Aggressions-Hemmung nach. Den Abschluß des empirischen Teils bilden erste Untersuchungen zur Validität. Sie tragen zum Teil noch den Charakter von Voruntersuchungen; soweit das der Fall ist, halten wir ihre Darstellung dennoch für gerechtfertigt, weil sie mit ihren Ergebnissen eine erste "Zwischenbilanz" zu dem definitiv nicht abschließbaren Validitätsproblem sinnvoll ergänzen.

Die zur Entwicklung unseres Testverfahrens erforderlichen Berechnungen wurden am Rechenzentrum der Universität des Saarlandes durchgeführt. Zum Teil wurde eine Anlage vom Typ TR 440 (TR), zum Teil eine Anlage vom Typ SIEMENS 7.561, BS 2000 (SI) benutzt. Standardmäßig werden unter den Ergebnistabellen und -abbildungen sowie im Text die o.g. Rechner mit den angegebenen Abkürzungen genannt und dahinter die benutzte Software mit ggfls. erforderlichen Spezifikationen aufgeführt.

Der Autor möchte an dieser Stelle Dank sagen für konstruktive Anregungen und gewährte Unterstützung bei der Durchführung der einzelnen Untersuchungen Herrn Prof. Dr. Tack, Herrn Prof. Dr. Kornadt und ganz besonders Herrn Dr. Zumkley.

I. THEORETISCHER TEIL

1. Einleitung

Im Gegensatz etwa zur Erforschung des Leistungshandelns hat die empirisch-psychologische Persönlichkeitsforschung aggressives Handeln bevorzugt unter einer sozialpsychologischen Perspektive untersucht (vgl. FESHBACH 1970; SELG 1974; BARON 1977; HECKHAUSEN 1980, S. 349 ff.; KORNADT 1980, 1981 a; PARKE & SLABY 1983). Durch die Pionierarbeit von DOLLARD et al. (1939) in Aufschwung gebracht, hat sie sich - mit zunehmendem Einfluß der sozialen Lerntheorie der Aggression (BERKOWITZ 1962, 1974; BANDURA & WALTERS 1963; BANDURA 1973) - vorwiegend der Differenzierung von Situationsfaktoren zugewandt, die mit Aggression einhergehen, und individuelle Unterschiede von Personen in diesem Zusammenhang vernachlässigt (HECKHAUSEN 1978; KORNADT & ZUMKLEY 1982, S. 310; ZUMKLEY 1984 a). Konsequent zu Ende gedacht, favorisiert der sozialpsychologische und lerntheoretische Bezugsrahmen eine systemtheoretische Betrachtungsweise, die dispositionelle Unterschiede der Aggressivität zwar nicht gänzlich ausschließt, ihnen im einzelnen jedoch eine nur untergeordnete Rolle beimißt (z.B. STRAUS 1974; MUMMENDEY, LINNEWEBER & LÖSCHPER 1984).

Ohne die Bedeutung situativer Determinanten, die für Aggression theoretisch als auch empirisch gut belegt sind (vgl. BARON 1977; PATTERSON 1979; HECKHAUSEN 1980, S. 349 ff.; PARKE & SLABY 1983), schmälern zu wollen, scheint u.E. jedoch ein Standpunkt sinnvoll, dem zufolge diese in Abhängigkeit von individuellen Merkmalen der

Person eine jeweils ganz unterschiedliche Bedeutung haben können und daher - bei objektiv gleichen Bedingungen der Situation - eine interindividuelle Varianz im Handeln zur Folge haben. Danach stellt eine einfache Wenn-dann-Beziehung zwischen Merkmalen der Situation und aggressivem Handeln, die die individuelle Wahrnehmung, Deutung und Bewertung der Situation nicht in Rechnung stellt, ein unzureichendes Modell dar, das die Bedingungsklärung aggressiven Handelns unnötigerweise einschränkt. Von ähnlich unzulänglichen Voraussetzungen waren bereits trieb- und instinkttheoretische Modelle im Vorfeld der lerntheoretisch begründeten Aggressionsforschung (ADLER, FREUD, McDOUGALL; vgl. KORNADT 1981 a) ausgegangen, indem sie den Verursachungskomplex für aggressives Handeln - ebenso einseitig - innerhalb der Person lokalisiert und dabei situativen Faktoren kaum Beachtung geschenkt hatten.

Die oben angedeutete interaktionistische Perspektive, die, von einer differenzierten, wechselseitigen In-Beziehung-Setzung von Merkmalen der Situation und Merkmalen der Person ausgehend, die Frage zu beantworten sucht, wie beide Variablengruppen (Situations- und Personvariablen) im Einzelfall zusammenwirken, ist jedoch erst in den letzten Jahren - insbesondere im Rahmen der Konzeptualisierung integrativer, motivationstheoretischer Modelle (OLWEUS 1978; KORNADT 1974, 1980, 1982 a, c, 1984) - angemessen berücksichtigt worden. Unter den empirischen Befunden, die für die Fruchtbarkeit dieser Betrachtungsweise sprechen, sei beispielhaft auf zwei neuere Arbeiten verwiesen. So konnten DODGE & FRAME (1982) bei Kindern und

ZUMKLEY (1984 a, b) bei Erwachsenen zeigen, wie individuelle Differenzen der Aggressivität Informationsverarbeitung und Verhalten in aggressionsrelevanten Situationen beeinflussen, d.h. situationale Variablen in ihrer Auswirkung auf aggressives Handeln modifizieren (vgl. dazu auch PARKE & SLABY 1983, S. 570 ff.).

Im Rahmen einer motivationstheoretischen und in dem o.g. Sinne interaktionistischen Betrachtungsweise der Aggression (KORNADT 1974, 1980, 1982 a, c, 1984) stellt sich die Aufgabe, individuelle Unterschiede in zwei Persönlichkeitsdispositionen, die dazu als handlungswirksam konzipiert wurden, im Aggressionsmotiv und in der Aggressionshemmung, meßbar zu machen. KORNADT (1982 a, b) hat ein Verfahren vorgestellt, das dieser Aufgabenstellung bei männlichen Jugendlichen und Erwachsenen dient und gerecht wird. Für Kinder liegt ein vergleichbares Verfahren noch nicht vor, was uns dazu bewogen hat, die Entwicklung eines solchen ins Auge zu fassen und zu versuchen.

Bevor wir auf diese Problemstellung näher eingehen, möchten wir zunächst in Kürze auf die durch Konstrukt- und Operationalisierungsvielfalt kontroverse theoretische Auseinandersetzung mit der menschlichen Aggression zu sprechen kommen und einige Verfahren vorstellen, die dazu entwickelt wurden, den Personfaktor "Aggressivität" zu erfassen. Wir beenden den ersten Teil dieser Arbeit dann mit der Darstellung der theoretischen und konstruktionstechnischen Rahmenvorstellungen, die die Entwicklung des hier vorgestellten Verfahrens geleitet haben. Über letztere werden wir im zweiten, empirischen Teil der Arbeit Bericht erstatten.

2. Aggression

Die Aggressionsforschung hat bis heute noch kein solides und allgemein anerkanntes theoretisches Fundament errichten können (KORNADT 1982 c, S. 91 f.). Neben einer Vielfalt theoretischer Konzeptionen äußert sich dieser Sachverhalt zunächst in einer nach wie vor uneinheitlichen Verwendungsweise des Begriffs "Aggression".

2.1 Begriff

Obwohl sich forschungsgeschichtlich dazu keine Parallele aufzeigen läßt, beginnt das Einstiegsproblem bei einer systematischen Behandlung von Aggression mit der Frage, wie sie definitorisch festgelegt und damit von anderen untersuchten Gegenstandsbereichen abgegrenzt werden soll.

In der Psychiatrie etwa sind mit "Aggression" schon ganz unterschiedliche Sachverhalte bezeichnet worden: ein innerer Zustand von Ärger, ein Impuls/Wunsch zur Destruktion, offene Akte von Zerstörung oder gar der Anspruch, Menschen und Dinge zu beherrschen (vgl. ZEGANS 1971). Ein Beispiel dafür aus neuerer Zeit liefert PARENS (1979). Dieser - meist tiefenpsychologisch orientierten - Verwendungsweise des Begriffs liegt kein erkennbares Bemühen um definitorische Klarheit zugrunde.

Anders dagegen schon bei physiologischen und ethologischen Untersuchungen von Aggression am Tier. Dort haben sich sogenannte

"topographische" Ansätze zur Begriffsbestimmung herausgebildet und bewährt. Sie gehen davon aus, daß durch genaue Beobachtung aggressionstypische Verlaufsgestalten des Verhaltens ermittelt werden können, die sich an das Vorhandensein spezifischer situativer Bedingungen knüpfen lassen. MOYER (1973) z.B. unterscheidet in dem o.g. Sinne "Beute-Aggression" ("predatory aggression") von "Furcht-induzierter Aggression" ("fear-induced aggression") und weiteren Unterarten, bei denen sich in einigen Fällen die neuralen und endokrinen Grundlagen überlappen. Für menschliche Aggression sind nach diesem Begriffsansatz noch keine brauchbaren Definitionsvorschläge formuliert worden (PARKE & SLABY 1983). ANGST (1980, S. 69) meint, ein spezifisch menschliches Feld aggressionsauslösender Situationen entstünde mit der Durchbrechung kultureller Normen.

Andere Autoren schenken den situativen und innerorganismischen Bedingungen, die Aggression vorangehen, besonderes Augenmerk. LORENZ (1963) sieht in Aggression einen spezifischen, im offenen Verhalten zur Entladung gebrachten Instinkt; das Todestrieb-Konzept der psychoanalytischen Theorie (FREUD; vgl. Abschnitt 2.2) geht von einer vergleichbaren Vorstellung aus. Wegen der Herstellung einer kausalen Beziehung zwischen Aggression und vorangehender Frustration und der genauen Umschreibung von Frustration in situationsorientierten Begriffen (Unterbrechung einer erwarteten Verhaltenssequenz und damit Verhinderung des Eintretens der Zielreaktion) läßt sich auch die Aggressionsdefinition von DOLLARD et

al. (1939) begründbar hier anführen. Ein weiterer Grund ist, daß sie die Antizipation eines spezifischen Handlungsergebnisses beinhaltet. Die Autoren der genannten Veröffentlichung verstehen unter Aggression "jede Verhaltenssequenz, deren Zielreaktion die Verletzung der Person ist, auf die sie gerichtet ist" (zit. n. KORNADT 1981 a, S. 24). Daneben räumen sie der Verletzungsintention eine für die Definition konstitutive Bedeutung ein. Eine der Schwierigkeiten bei Definitionen dieser Art ergibt sich dadurch, daß man sich auf erschlossene Sachverhalte stützt (Instinkt, Trieb, Intention u.ä.), die einer zuverlässigen und gültigen Feststellung erhebliche Schwierigkeiten bereiten, sodaß man es zum Teil erst garnicht versucht hat (LORENZ, FREUD). Zudem verwenden letztgenannte Autoren "Aggression" begriffslogisch unsauber (zirkulär).

Mit aus o.g. Gründen hat BUSS (1961) sich gegen das intentionale Moment der Definition von DOLLARD et al. (1939) gewandt und mit einem eigenen Vorschlag ("a response that delivers noxious stimuli to another organism"; ebd., S. 1) versucht, dem Wissenschaftsideal eines streng behavioristischen Operationalismus gerecht zu werden. Ein Problem dabei ist, daß Selbstaggressionen mit dem BUSS´schen Vorschlag ausgeklammert werden; ein weiteres, daß nicht immer ohne weiteres festgestellt werden kann, ob überhaupt eine Verletzung vorliegt. Da der die Handlung leitenden Motivation nicht Rechnung getragen wird, sind Aggressionen, die ihr Handlungsziel verfehlen, definitiv keine, das tägliche Handwerk eines Dentisten dagegen besteht im wesentlichen aus Aggression. Wie schwer es ist, die

Intention auszuklammern, zeigt uns BUSS mit einem Beispiel selbst: Ghandi's Hungerstreik in Indien gegen die Briten bezeichnet er als eine "passive, indirekte Aggression" (BUSS 1961, S. 9).

Letztgenanntes Beispiel verweist deutlich darauf, daß die einer Handlung zugrundeliegende Verletzungsabsicht das zentrale Bestimmungsstück für Aggression darstellt, die nach FESHBACH (1964; zit. n. KORNADT 1981 a, S. 34) "feindseliger" Art ist, wenn die Verletzung selbst das eigentliche (primäre) Handlungsziel darstellt, "instrumenteller" Art dagegen, wenn letztere in Kauf genommen bzw. herbeigeführt wird, um damit ein anderes Handlungsziel zu erreichen. Das Beispiel verweist aber noch auf einen anderen bedeutsamen Aspekt. Handlungsergebnis ist die körperliche Schwächung Ghandi's, die Folge erwartete Betroffenheit der Briten und dadurch motiviertes Einlenken. Mithin ist nicht das Handlungsergebnis selbst, sondern der Charakter der intendierten Handlungsergebnisfolge (hier: Betroffenheit) entscheidend. Da objektiv jedoch nur das Einlenken zu beobachten war, mag hieran deutlich werden, daß die subjektive Bestimmungsleistung, Aggression läge vor, leicht durch ihre Interpretationsanteile überdehnt werden kann. Der "social judgement approach" (PARKE & SLABY 1983, S. 550) betrachtet Aggression gerade unter dem Gesichtspunkt einer soziokulturellen Bestimmungsleistung, in der gruppenspezifische Normen und Standards eine wesentliche Rolle spielen. Danach gibt es eigentlich nichts Aggressionsspezifisches mehr als objektiven Sachverhalt.

Als Resümee läßt sich folgendes festhalten: Man ist in dem Bemühen um begriffliche Klarheit bis heute zwar zu keiner allgemein akzeptierten, einheitlich-definitorischen Festlegung gekommen, was menschliche Aggression anbelangt (vgl. JÜTTEMANN 1982; MUMMENDEY 1982; PARKE & SLABY 1983); jedoch hat sich dabei eine verschärfte und m.E. gewinnbringende Betrachtungsweise des Gegenstandes herausgebildet. Andererseits hat der unterschiedliche Gebrauch des Begriffs "Aggression" in der einschlägigen Forschung erheblich mit dazu beigetragen, die intra- und interdisziplinäre Vergleichbarkeit ihrer Ergebnisse zu erschweren (vgl. KORNADT 1981 a). So ist z.B. der in vielen lerntheoretischen Experimenten zur menschlichen Aggression fehlende Nachweis einer vorliegenden Schädigungsabsicht bzw. allgemein der die Handlung leitenden Motivation zum Ansatzpunkt der Kritik geworden, an deren genereller Validität für (feindselige) Aggression zu zweifeln (vgl. MICHAELIS 1976; ZUMKLEY 1980; MUMMENDEY 1982). Diese und weitere Erschwernisse einer stringenten definitorischen Festlegung mögen mit zu dem Vorschlag beigetragen haben, den Begriff überhaupt als solchen aufzugeben (vgl. KORNADT 1982 c). Als eine Art kleinsten gemeinsamen Nenner kann man vorerst die Vorstellung betrachten, mit "Aggression" ein Verhalten zu bezeichnen, das absichtlich herbeigeführt wird, um damit einer anderen Person bzw. anderen Personen zu schaden oder Unrecht zuzufügen (vgl. PARKE & SLABY 1983).

2.2 Theoretische Konzepte

Seit der vorsokratischen Naturphilosophie sind über die Natur der Aggression ganz unterschiedliche Vorstellungen formuliert und bekannt geworden. Es blieb jedoch dem 20. Jhd. n. Chr. vorbehalten, das Phänomen im Sinne eines angreifenden, Schmerz und Leid zufügenden Verhaltens und/oder einer zugehörigen Persönlichkeitsdisposition (Aggressivität) zu betrachten.

Mit einer der ersten, die sich mit dieser Aufgabe befaßten, war Sigmund FREUD (vgl. FROMM 1974; KORNADT 1981 a). Er stellte im Laufe der Jahre drei verschiedene Aggressionstheorien (-Modelle) auf, aus denen die nachfolgende Forschung zum Teil entscheidende Anregungen bezog.

Nach dem zuerst formulierten Modell betrachtete FREUD die Aggression als eine besondere Komponente libidinöser (sexueller) Triebe, die sich schon sehr früh, etwa in dem Bedürfnis nach Einverleibung des Sexualobjektes, zeige. Etwas später, in der ödipalen Phase, sei die Liebe eines Kindes zum gegengeschlechtlichen Elternteil behaftet mit Todeswünschen gegenüber dem rivalisierenden gleichgeschlechtlichen.

Ab 1915 gab FREUD jedoch diese Vorstellungen auf, um nun, in seinem zweiten Modell, das Reaktive in der Aggression zum Ansatzpunkt eines neuen Konzeptes zu erheben, das er selbst jedoch nur als eine Übergangslösung betrachtete. Aggression stellt demzufolge eine soziale Reaktion auf Streß und umweltinduzierte Frustration dar,

die das ICH in Gang setzt, um die Befriedigung instinktiver Bedürfnisse zu gewährleisten. Das o.g. Beispiel war demzufolge nun so zu deuten, daß der gleichgeschlechtliche Elternteil dem Kinde die Liebe zum gegengeschlechtlichen behinderte (versagte), worauf dieses mit Aggression reagierte.

Das dritte Modell schließlich entstand im Zusammenhang mit FREUDs Revision seiner gesamten Trieblehre. Ab 1920 postulierte er neben dem Eros, in dem er nun die libidinösen und die ICH-Triebe zusammenfaßt, einen zweiten fundamentalen Trieb, den Todestrieb. Darunter verstand er eine dem Eros entgegenwirkende Kraft mit der Tendenz zur totalen physischen (und für FREUD damit auch psychischen) Entropie. Sie sei primär auf die Zerstörung des Individuums selbst gerichtet. Im Zusammenspiel mit dem konstruktiven Eros ergebe sich jedoch eine Umleitung der Todesenergie nach außen, wo sie dann als Aggression in Erscheinung trete.

In kritischer Würdigung des FREUDschen Gedankenguts läßt sich folgendes feststellen: Das erste Modell zeugt davon, daß FREUD Aggression als eigenständiges Phänomen gesehen und versucht hat, es in seinem theoretischen Gebäude unterzubringen. Das zweite Denkmodell hat die nachfolgende Forschung entscheidend angeregt und in Fluß gebracht, obgleich FREUD selbst ihm skeptisch gegenüberstand. FREUDs Skeptizismus war berechtigt, denn er hatte damit nichts grundlegend Neues formuliert. Mit THOMAS VON AQUIN (1225 - 1274) hatte die Philosophie der Hochscholastik bereits die Grundannahmen der modernen Reaktivitätstheorie vorweggenommen (vgl. WINDELBAND

& HEIMSOETH 1957). Neu und bedeutsam wurden jedoch demgegenüber die von FREUD in dieser Phase seines theoretischen Schaffens entwickelten Vorstellungen zu verschiedenen Abwehrmechanismen bei Aggressionskonflikten. Das dritte Modell schließlich beruht auf einem metaphysischen Konzept FREUDs zur Deutung der tieferen Ursachen der Wirklichkeit schlechthin und beinhaltet nicht etwa die Annahme einer physiologisch-definierbaren Triebquelle der Aggression (vgl. FROMM 1974). Es hat - unverkennbar Anleihen bei der vorsokratischen Naturphilosophie (EMPEDOKLES; vgl. WINDELBAND & HEIMSOETH 1957) aufweisend - bis heute überlebt, in moderner Abwandlung etwa bei MITSCHERLICH (1969).

In Anlehnung an FREUDs zweites Modell veröffentlichten DOLLARD et al. (1939) den ersten lerntheoretisch orientierten Erklärungsansatz zur Aggression, der als "Frustrations-Aggressions-Theorie" (F-A-Theorie) bekannt wurde und die Forschung bis heute beeinflußt (vgl. HILKE & KEMPF 1982). In ihrer ursprünglichen Fassung nahm die Theorie einen ursächlichen Zusammenhang zwischen Frustration und Aggression als einer spezifischen Folgereaktion darauf an. Erstere führe immer zu letzterer, und aus dem Vorliegen letzterer ließe sich immer auf das Vorangehen ersterer schließen. Durch die Erwartung von Mißerfolg bei und Strafe für Aggression z.B. kann Aggression daneben gehemmt werden. Durch besondere Mechanismen der Verarbeitung des daraus resultierenden Konflikts (z.B. Verschiebung) wird Aggression nach außen dann abgewandelt zum Ausdruck gebracht (indirekter Formenkreis) oder richtet sich gar nach innen, gegen das

eigene Selbst. Gleichermaßen sozialpädagogisch als auch psychotherapeutisch relevant war eine als "Katharsis-Hypothese" zusätzlich formulierte Annahme, nämlich "daß das Auftreten einer Aggressionshandlung die Anregung zur Aggression reduziert" (zit. n. KORNADT 1981 a, S. 26).
Zwei Jahre später (MILLER 1941) hat die Theorie eine wichtige Revision erfahren. Man räumte nun ein, daß auch nicht-aggressive Reaktionen Folge von Frustration sein können; das Auftreten von Aggression selbst setze jedoch immer Frustration voraus.

Eine besondere Bedeutung - über den angesprochenen Phänomenbereich hinaus - erlangte die F-A-Theorie vor allem deshalb, weil sie erstmals psychoanalytisches Gedankengut in einer empirisch nachprüfbaren Form vorgelegt hatte.

Bislang am einflußreichsten waren jedoch die in der Folge formulierten lerntheoretischen Konzepte zur Aggression, die zum Teil in unmittelbarer Auseinandersetzung mit der F-A-Theorie entwickelt wurden. An dieser Stelle sind vor allem die Theorien von BERKOWITZ und BANDURA zu nennen.

BERKOWITZ (1962) nahm - zunächst noch in enger Anlehnung an die F-A-Theorie - an, daß durch Frustration Ärger als mediatorische Antriebskomponente für Aggression resultiert (angeborener Anteil der als universell gedachten Verknüpfung von Aggression mit Frustration), die jedoch erst im Zusammenwirken mit richtungsweisenden situativen Hinweisreizen (gelernter Anteil der o.g. Verknüpfung)

spezifisch verhaltensrelevant wird. Dieses Modell einer auf Hinweisreize klassisch konditionierten Reaktion sollte vor allem die impulsiven Erscheinungsformen von Aggression erklären. Ab 1974 trug BERKOWITZ dann auch den Bekräftigungsfolgen aggressiver Handlungen, die zuvor schon von Theoretikern wie SEARS (1958) oder BUSS (1961) hervorgehoben worden waren, explizit Rechnung. Darüber hinaus nahm er an, daß Hinweisreize eine aggressionssteigernde Wirkung haben können.

Kritisch anzumerken bei BERKOWITZ ist, daß er die Bedingungen, unter denen aggressionsrelevante Hinweisreize entstehen, nicht genau zu spezifizieren vermochte und ebensowenig den dazu erforderlichen unkonditionierten Reiz (vgl. ZUMKLEY 1983).

Nicht mehr den Druck aversiver Zustände ("push-Konzept"), sondern den Anreizwert erwarteter Handlungsergebnisfolgen ("pull-Konzept") rückte die soziale Lerntheorie der Aggression (BANDURA 1973) in den Vordergrund ihres Bezugsrahmens. Aufgrund einer breit angelegten experimentellen Forschung bereicherte sie die theoretische Diskussion um Aggression durch Kenntnisse zur Wirkung imitativer Prozesse. BANDURA & WALTERS (1963) konnten zunächst zeigen, daß auch stellvertretende Erfahrung in der Form von Beobachtung aggressiver Modelle zum Erwerb aggressiver Handlungsmuster führen kann. Deren Ausführung hatte die Theorie zunächst von den ebenso gelernten und anschließend erwarteten Handlungsergebnisfolgen abhängig gemacht. Damit war eine klare Alternative zur F-A-theoretischen Deutung der Ursachen des

Auftretens von Aggression formuliert. Den Ärger betrachtet BANDURA (gegenüber BERKOWITZ) nicht mehr als wesentlich für Aggression, sondern als eine spezifisch nachträglich gedeutete, primär unspezifische Erregung.
In der Folge bemühte sich BANDURA dann um eine sehr dezidierte Erklärung der zwischen Beobachtung und Ausführung von Aggression gelegenen moderierenden Zwischenprozesse. In diesem Zusammenhang kam er zu der Auffassung, daß neben fremdverabreichten Bekräftigungen auch solche der Selbstbekräftigung eine wichtige Rolle spielen. Über ein sog. "self-monitoring-reinforcement-system" können leicht gesellschaftliche Normen und Verhaltensstandards, die als persönlich verbindlich erlebt und angesehen werden, in Handlungsentscheidungs-Prozesse einfließen. Daher kann unter gleichen situativen Umständen auch leicht anders reagiert werden, als mit Aggression (z.B. mit Abhängigkeit, Leistung, Rückzug etc.). Damit rückt BANDURA deutlich den selbstregulativen Charakter aggressiven Handelns in den Blickpunkt und überschreitet die assoziative Wechselwirkungsmechanik zwischen situativen Hinweisreizen und internalen Triebzuständen bzw. Entladungsdrängen endgültig.

Die Prozesse der kognitiven Aufarbeitung und Bewertung des Verhaltens und der Situation, die sehr viel umfassender als einfache, intervenierende Regelgrößen angelegt sind, wurden jedoch, obgleich sie Ansatzpunkte dazu bieten, nicht zur Ableitung individueller Differenzen der Aggressivität herangezogen und für die Forschung nutzbar gemacht. Das lag wohl mit auch daran, daß diese Theorie

Aggression in einen so weit gespannten Rahmen integriert hatte, daß sie in einer allgemeinen Theorie sozialen Handelns quasi "unterging". Aggression wird zum Verhalten schlechthin; es gibt an ihr nichts Spezifisches mehr (vgl. KORNADT 1982 c). Kritisch anzumerken bei BANDURA ist auch, daß über das Wie des Zusammenwirkens der vielen hypothetischen Prozesse kaum etwas Verbindliches ausgesagt wurde.

Weitere Konzepte, die in die psychologische Diskussion um Aggression Eingang gefunden haben, wurden in Ethologie und Neurophysiologie entwickelt.

Konrad LORENZ (1963) repräsentiert zwar nicht den ethologischen Standpunkt (vgl. ANGST 1980), ist aber mit seiner Publikation am meisten bekannt geworden. "Intraspezifische Aggression" nennt LORENZ den auf den Artgenossen gerichteten Kampftrieb von Mensch und Tier und hebt davon "interspezifische Aggression" in der Bedeutungsumschreibung "zwischenartliche Kämpfe" ab. Er setzt damit die neodarwinistische Theorie MacLEANs fort, der Aggression zu den vier überlebensnotwendigen Grundtrieben einer Art rechnet (vgl. MacLEAN 1958). Demzufolge hat LORENZ vor allem den Anpassungswert der Aggression im Evolutionsprozeß der Arten herausgestellt als auch ihre kommunikative und systemregelnde Funktion (etwa bei Etablierung und Stabilisierung einer sozialen Struktur) betont. Sein Schüler LEYHAUSEN (1967, S. 283) hat letztgenannte Annahme später dahingehend überdehnt, daß er im Antagonismus von Angst und

Aggression überhaupt die Grundlage aller sozialen und kulturellen Ordnung vermutet.
Im übrigen vertritt LORENZ ein sog. "Dampfkesselmodell". Ein hypothetischer neuraler Behälter wird zyklisch mit aggressionsspezifischer Energie ge- und nach Überschreiten eines kritischen Regelwertes entladen, bei Abwesenheit situativer Auslöserreize in eine sog. "Leerlaufhandlung".

LORENZens evolutionsbiologisches Modell wurde in Fachkreisen weitgehend als "Ausrutscher" des Nobelpreisträgers auf einem Gebiet verstanden, auf dem er es versäumt hatte, sich zuvor fachkundig zu machen. Dem persönlichen Engagement in einer sozial brennenden Frage trug jedoch die große Publizität seiner Veröffentlichung Rechnung. Zur ausführlichen Kritik vgl. z.B. FROMM (1974, S. 16 ff.).

MOYER (1971 a, b, 1973) hat die Triebhydraulik des LORENZ-Modells kritisiert und ist mit einem multifaktoriellen, neurophysiologischen Ansatz von sehr viel differenzierteren Annahmen ausgegangen. Spezifische aggressionsrelevante Areale im Gehirn leiten nach seinen Vorstellungen bei Überschreiten eines kritischen Verhältniswertes zu analog konzipierten neuralen Hemmungsarealen im Zusammenwirken mit situationstypischen Reizen spezielle Unterarten von Aggression (Beute-Aggression, Rivalen-Aggression etc.) in die Wege. Daneben nennt MOYER eine Reihe weiterer Faktoren (z.B. hormonelle), anerkennt die Bedeutung instrumentellen Lernens und hält Aggression insgesamt auf einer ähnlichen Grundlage wie etwa

Nahrungsaufnahme oder Paarungsverhalten für erklärbar.

Neben einer Reihe logischer Schwierigkeiten im Detail, etwa was die Abgrenzung unterschiedlicher Arten von Aggression oder die Beweisführung für das Vorliegen spezifischer Aggressionsimpulse angeht, kann MOYER's Konzept für menschlich aggressives Handeln bislang nicht als bestätigt angesehen werden, zumal er sich selbst widersprüchlich dazu äußert, ob es nach der empirischen Befundlage überhaupt auf den Menschen übertragbar ist.

Wichtig bei all den vorgestellten (und nicht vorgestellten) Konzepten ist, daß keines davon Aggression - insbesondere menschliche - in ihrer Komplexität und phänomenalen Vielfalt auch nur annähernd umfaßt. Die meisten umschreiben und erklären wichtige Teilaspekte, sind hinsichtlich dieser auch empirisch fundiert und beziehen von daher ihre Legitimation und partielle Gültigkeit. Darauf hat insbesondere MICHAELIS (1976) hingewiesen und versucht, die unterschiedlichen Ansätze in ein kybernetisches Prozeßmodell zu integrieren.

Neuere Entwicklungen in der Theorienbildung (vgl. GEEN & O'NEAL 1976; HILKE & KEMPF 1982; PARKE & SLABY 1983) betonen immer stärker kognitive Prozesse (z.B. Attribution), den handlungssteuernden Charakter von Normen der jeweiligen Sozietät und systembezogene Regel-Aspekte der vorhandenen Situation. Das Konzept von DODGE (1982) betrachtet Aggression aus der Perspektive eines

Handlungs-Entscheidungs-Prozesses, in den die o.g. Aspekte in einer spezifischen Art und Weise eingeflochten sind (vgl. PARKE & SLABY 1983, S. 556 ff.).

2.3 Zusammenfassung

Nach wie vor ist unklar, was unter dem Begriff "Aggression" sinnvollerweise verstanden werden soll. Bemühungen um eine Präzision seiner begrifflichen Fassung einerseits steht der Vorschlag andererseits gegenüber, ihn als solchen aufzugeben (vgl. KORNADT 1982 c). Minimaler Konsens dürfte mit der Vorstellung einhergehen, damit ein Verhalten zu bezeichnen, das absichtlich herbeigeführt wird, um damit einer anderen Person bzw. anderen Personen Schaden zuzufügen oder Unrecht zu tun (vgl. PARKE & SLABY 1983).

Überblickt man die Theorienlandschaft zum o.g. Gegenstandsbereich, zeichnet sich ein ähnliches Bild ab. Die Triebtheorie der Aggression spannt einen weiten Bogen von metaphysischen, naturphilosophischen (FREUD) bis zu neurophysiologischen, naturwissenschaftlichen Denkansätzen (MOYER). Ihre Protagonisten halten sie i.d. Regel für ein wertfreies und objektives Paradigma. Paradoxerweise stellt sie jedoch ein streng deterministisches Denkmodell dar, zu dessen Verabsolutierung ihm die empirische Grundlage fehlt und dessen naturphilosophische Begründung zumindest als fraglich, wenn nicht als überholt gelten muß (vgl. WEIZSÄCKER 1982).

Die lerntheoretischen und kognitionstheoretischen Erklärungsansätze

stellen ein nicht minder breites Spektrum dar. Sie reichen von relativ orthodoxen, klassisch-behavioristischen Modellen (SEARS) aus der Mitte dieses Jahrhunderts bis zu neuzeitlichen, dem "Computerzeitalter" und der "kybernetischen Gesellschaft" entsprechenden Informationsverarbeitungsmodellen (DODGE). Erinnern erstere an die Blütezeit des angelsächsischen Empirismus, so sind letztere ein "Kind unserer Zeit". Auch sie bringen unmittelbar ein bestimmtes Verständnis von Wertfreiheit und "Wissenschaftlichkeit" zum Ausdruck, ohne es jedoch unmittelbar zu hinterfragen.

Einem neueren Trend zufolge wächst die Einsicht, von einer jeweils partiellen Gültigkeit der einzelnen Modellansätze auszugehen und die damit belegten Befunde in einem integrativen Modell zusammenzufassen (vgl. MICHAELIS 1976; KORNADT 1982 c). Diese Aufgabe liegt jedoch im wesentlichen noch vor uns.

Von eminent praktischer als auch sozialpolitischer und strafrechtlicher Bedeutung sind jedoch die aus den einzelnen Modellansätzen ableitbaren Schlußfolgerungen hinsichtlich der Bedingungsklärung und darauf beruhenden Bewertung von (insbesondere feindseliger) Aggression (vgl. z.B. FÜLLGRABE 1983) sowie der darauf beruhenden Konsequenzen pädagogischen Handelns (KORNADT 1982 c; ZUMKLEY 1983).

3. Diagnostik der Aggression

SELG (1974) hat - nach einer gründlichen Durchsicht vorhandener Verfahren zur Erfassung von Aggression und Aggressivität - resignierend vom "Elend der Aggressivitätsdiagnostik" gesprochen, die bis dahin nur "vorgestaltartige Ergebnisse" geliefert habe. Dieses Bild läßt sich auch heute noch zeichnen. Es wäre daher von fraglichem Nutzen, auf eine systematische Darstellung und Behandlung dieses Themas an dieser Stelle noch einmal einzugehen. Wir wollen uns statt dessen auf die Darstellung einiger ausgewählter Erhebungsverfahren beschränken und in diesem Zusammenhang auf einige ihrer Probleme verweisen. Damit soll einerseits verdeutlicht werden, wie Diagnostik der Aggression konkret betrieben worden ist und wird, andererseits dem Leser eine (grobe) Orientierungsgrundlage geboten werden, die ihm die testpsychologische Einordnung des in dieser Arbeit vorgestellten Verfahrens erleichtert. Wir hoffen damit aber auch zugleich zeigen zu können, vor einem wie vielfältigen, ungesicherten und problematischen Hintergrund wir unsere eigene Arbeit aufgenommen haben.
Zum besseren Verständnis der einzelnen Verfahren wollen wir zuvor auf unterschiedliche Auffassungen hinsichtlich ihrer diagnostischen Aufgabenstellung zu sprechen kommen, die mit deren Konzeption unmittelbar zum Ausdruck gebracht werden. Daneben kurz auf unterschiedliche Meßmodelle und testmethodische Konzepte Bezug nehmen, nach denen die Verfahren im einzelnen gestaltet worden sind sowie auf methodisch bedingte Vor- und Nachteile.

3.1 Perspektiven diagnostischer Aufgabenstellung

Der Artikel von GOLDFRIED & KENT (1972) beschreibt eine Umorientierung in der Auffassung von der Aufgabenstellung psychodiagnostischer Verfahren. Mit diesem Wandel ging eine Änderung in der Konzeption persönlichkeitstheoretischer Modelle einher.

In den traditionellen, personzentrierten Modellen war man grob gesagt der Auffassung, daß offenem Verhalten eine (hypothetisch erschlossene) Persönlichkeitsstruktur zugrundeliegt, die mit diesem in einem funktionalen Zusammenhang steht. Für diese hypothetisch geforderte Struktur hatte man sich unterschiedliche Theorien ausgedacht, die einerseits mehr den funktionalen (psychodynamische Persönlichkeitsmodelle), andererseits mehr den strukturellen Aspekt betonten (trait-Modelle). Diagnostisches Ziel in den genannten Modellen war das richtige und genaue Abbild funktionaler und struktureller Aspekte der Persönlichkeit. Man glaubte, mit dieser Information voraussagen zu können, was eine bestimmte Person in einer bestimmten Situation tun würde. Die Versuche, diesen Denkansatz empirisch zu stützen, waren bis dahin jedoch nicht sonderlich überzeugend ausgefallen und hatten MISCHEL (1968) zu einer Kritik veranlaßt, die in der Folge die sog. "Situationismus-Debatte" auslöste (vgl. z.B. MISCHEL 1973; BOWERS 1973; BEM & ALLEN 1974; WACHTEL 1977).

In den nun favorisierten, situationszentrierten Modellen nahm man an, daß offenes Verhalten besser (als bislang) vorausgesagt und erklärt werden könne, wenn man statt dessen über Information verfügte,

die sich auf die funktionalen Beziehungen zwischen diesem und Merkmalen der Situation, in der es stattfand, bezog. Wichtig war nun statt dessen - ganz pragmatisch - die "nützliche Prognose", die man ohne den Umweg über erschlossene Persönlichkeitsmerkmale erstellen wollte, und diagnostisches Ziel demzufolge eine möglichst genaue Erfassung des situativen Kontexts und seiner charakteristischen und relevanten Besonderheiten, in dem die Person ein bestimmtes (auffälliges, von der Norm abweichendes) Verhalten zeigte. Beheimatet war dieser "verhaltensdiagnostische Ansatz" in den lerntheoretisch begründeten Persönlichkeitsmodellen (vgl. ANGLEITNER 1980, S. 154 ff.).

Das oben in aller Kürze skizzierte personzentrierte Vorgehen einerseits und das situationszentrierte Vorgehen andererseits hat auch in der Entwicklung von Testverfahren zur Diagnostik von Aggression seine Spuren hinterlassen. Ältere, personzentrierte Verfahren haben sich relativ wenig um den situativen Kontext von Aggression gekümmert und zum Teil ausschließlich Charakteristika im Innern der Person zu erfassen versucht (z.B. ELIZUR 1949), wohingegen neuere Entwicklungen eine zum Teil gerade umgekehrte Schwerpunktsetzung in der Gestaltung aufweisen (z.B. PETERMANN & PETERMANN 1980). Natürlich gibt es viele Zwischenformen von Verfahren, die zur Erfassung mehr oder weniger genau umschriebener Persönlichkeitsfaktoren entwickelt sind und ebenfalls mehr oder weniger eine (plausible) situative Orientierung erkennen lassen (z.B. DUHM & HANSEN 1957).

3.2 Meßmodelle und testmethodische Konzepte

Beginnen wir der Einfachheit halber mit Verfahren, die zur Erfassung eines trait-ähnlich gedachten Personfaktors "Aggressivität" entwickelt wurden. Diese Verfahren fassen Aggression als eine linear-monotone Funktion dieses Personfaktors auf. Je höher der Rang eines Individuums im Hinblick auf die mediatorische Variable "Aggressivität" ist, umso mehr (häufiger und/oder intensiver) wird es Aggression über verschiedene Situationen hinweg auch zeigen. Bei der Entwicklung eines derartigen Tests ist man an Verhaltens- und Befindlichkeitsmerkmalen einer Person interessiert, und zwar an solchen, von denen man glaubt, daß sie für die o.g. Persönlichkeitsvariable typisch sind. Ein Proband kann z.B. befragt werden, wie häufig er verärgert ist, aber auch, was er von einem Western hält.

Methode der Wahl ist hier der Persönlichkeitsfragebogen. Die danach entwickelten Verfahren rubriziert BRICKENKAMP (1975) unter den sog. "psychometrischen Persönlichkeitstests", was bereits auf ihre Vorteile verweist. Sie sind in der Regel durchführungs- und auswertungsobjektiv, und für ihre dimensionsanalytisch untersuchten Meßwerte sind ausreichende bis gute Meßgenauigkeiten nachgewiesen. Der Nachteil dieser Verfahren besteht zunächst darin, daß sie vergleichsweise leicht zu durchschauen sind. Es ist daher mit durch ein kritisches Bewußtsein kontrollierten, systematisch verfälschten Testreaktionen zu rechnen, die für das eigene Verhalten in den fraglichen Situationen keine Verbindlichkeit mehr besitzen. Daneben ist

generell fraglich, ob die das Verhalten entscheidend beeinflussenden Faktoren der Persönlichkeit unmittelbar ins Bewußtsein gelangen und dort einer direkten Befragung zugänglich sind. Jedenfalls geht man in der nächstdargestellten Gruppe von Testverfahren davon aus, daß sie im Verborgenen wirken, sich einer Bewußtwerdung und damit auch direkten Befragung entziehen.

Nach psychodynamischen Modellen konzipierte Aggressionstests verstehen demgegenüber zunächst einmal die funktionale Beziehung zwischen Personfaktor und offenem Verhalten als eine nicht-monotone, kurvilineare Funktion (vgl. MAGNUSSON & ENDLER 1976). Sie gehen davon aus, daß ab einem bestimmten, kritischen Punkt Hemmungsmechanismen in der Person ihre Wirkung entfalten, sodaß sich der zunächst monoton-steigende Trend umkehrt, sobald dieser Punkt überschritten wird. Nach dem Konflikt-Modell von MILLER können diese konflikttheoretisch orientierten Verfahrenskonzepte aus dem Zusammenwirken der antagonistischen Komponenten zusätzlich erklären, wie es - infolge überstarker Hemmung gegenüber dem (primären) Aggressionsziel - zu einer Verlagerung der Aggression gegen (unbeteiligte) Dritte oder Sachen kommen kann (vgl. z.B. CHRISTOPH-LEMKE 1972, S. 45 f.). Situative Faktoren werden auch hier nur unzureichend berücksichtigt; implizit (gegenüber den am trait-Modell orientierten Verfahren) jedoch insofern, als unter ungünstigen (inneren als auch äußeren) Bedingungen Aggression auch gehemmt werden kann und somit verlagert wird oder unterbleibt.

Bedeutsam für die am psychodynamischen Modell orientierten Verfahren

wurde die Entwicklung einer Phantasiegeschichten-Technik durch MURRAY (1943), bei der die Person gebeten wird, zu einer Reihe halbstrukturierter Bilder, die Ausschnitte aus verschiedenen sozialen Situationen des Menschen darstellen, phantasievolle Geschichten zu erfinden. In deren Gestaltung greifen dann, so die Annahme, unbewußt überdauernde und durch die Bildinhalte angeregte Motive ein, womit sie der Erfassung zugänglich werden. In den Geschichten suchte man zunächst nur Hinweise auf Aggression, später dann auch auf solche, die auf deren Hemmung verwiesen (z.B. MUSSEN & NAYLOR 1954; PURCELL 1956; LESSER 1958 a).

Nach BRICKENKAMP (1975) gehören die o.g. Tests zu den von ihm so genannten "Persönlichkeits-Entfaltungsverfahren", ein Begriff, der ihre eigentliche Stärke hervorhebt. Ohne daß dem Probanden die diagnostische Zielrichtung bewußt wird, kann er das Charakteristische seiner Persönlichkeit frei und ungezwungen in das Testverhalten einbringen, worauf sich die Validität dieser Verfahren stützt. Kennzeichnend ist auch die damit verbundene, mehr ganzheitlich und dynamisch orientierte Auswertung des Testverhaltens gegenüber der mehr elementaristisch-additiven der Fragebogenverfahren, was meist intuitiv als ein Beleg dafür angesehen wurde, daß die Ergebnisse valide seien. Unter einer psychometrischen Perspektive war jedoch letzteres gerade die Schwachstelle dieser Gruppe von Tests, die sich einer Item-Selektions-orientierten Entwicklung meßgenauer und dimensionsanalytisch geprüfter Verfahren in den Weg stellte.

Im verhaltensdiagnostischen Bezugsrahmen schließlich geht man, wie bereits angedeutet, davon aus, daß offenes Verhalten in einer direkten, funktionalen Beziehung mit dem spezifischen situativen Kontext steht, in dem es jeweils auftritt. Demzufolge attestiert man situationsbezogen-erhobenen Verhaltensstichproben die besten Prädiktorqualitäten für zukünftiges Verhalten in vergleichbaren Situationen. Man erhebt diese Stichproben mittels einer sog. "Verhaltensdiagnose". Zur Erfassung von Aggression etwa codiert ein in der Regel geschulter Beobachter nach einem Zeitplan und vorher festgelegten Verhaltenskriterien diverse ihrer phänomenologisch differenzierbaren Unterarten (verbale, körperliche, impulsive etc.) in lebensnahen Situationsausschnitten direkt (z.B. MEES & FIEGUTH 1977) oder schätzt deren Ausprägungsgrad mittels eines entsprechenden Rating-Verfahrens nachträglich ein (z.B. PETERMANN & PETERMANN 1978). Eine neuere Entwicklung "simuliert" Situationsausschnitte und Verhalten(smöglichkeiten darin) mittels vorgegebener Bild-Text-Kombinationen in einem nach Fragebogenprinzipien gestalteten Test (PETERMANN & PETERMANN 1980).

Als Hauptvorteil verhaltensdiagnostisch fundierter Verfahren betrachtet man die unmittelbar sich anbietenden Möglichkeiten zur Ableitung der sog. "Therapiezielmatrix" im Rahmen einer (ebenfalls lerntheoretisch begründeten) psychotherapeutischen Intervention, die dann vor allem aus Maßnahmen zur Veränderung des situativen Kontexts besteht (vgl. MEES 1974, 1982; PETERMANN & PETERMANN 1978; ZUMKLEY 1983). Demgegenüber ist die therapeutische Umsetzung von

Information personzentrierter Diagnostika weniger direkt und konkret und weist statt dessen mehr auf erforderliche Änderungen in der Persönlichkeitsstruktur des Klienten hin (vgl. DAVISON & NEALE 1984, S. 82 ff.), z.B. im Bereich aggressionsregulierender Normen. Ein Nachteil dieser Verfahren besteht darin, daß sie bislang noch nicht dort eingesetzt werden können, wo Aggression verhüllt und in sublimer Form sich äußert; ein weiterer darin, daß sie zum Aufspüren aggressionsrelevanter, kognitiv-regulativer Prozesse im Innern der Person (wohl prinzipiell) ungeeignet sind.

Kurz zusammenfassend kann man also zu der folgenden Feststellung gelangen: Die persönlichkeitstheoretisch orientierten Aggressivitätsmeßverfahren beruhen in der Regel auf einem "rationalen", die verhaltenstheoretisch orientierten Meßverfahren auf einem "empirischen Testkonzept" (vgl. LIENERT 1969, S. 53 f.). An ersteres ist ein mehr diagnostischer, an letzteres ein mehr prognostischer Verwendungszweck gebunden. Verfahren, die beide Schwerpunktsetzungen in der Gestaltung konstruktiv miteinander verbinden, gibt es vorläufig noch nicht. Bis dahin stehen beide Gruppen von Verfahren noch relativ unverbunden nebeneinander, und es nicht auszuschließen, daß damit unterschiedliche Gültigkeitsbereiche korrespondieren (vgl. ZUMKLEY 1985).

3.3 Darstellung einiger ausgewählter Erhebungsverfahren

Unter den bereits genannten Zielsetzungen beschränken wir uns auf die Vorstellung einiger weniger Testverfahren und Erhebungstechniken, für deren Auswahl folgende Gesichtspunkte maßgeblich gewesen sind: Zunächst sollten möglichst unterschiedliche theoretische und methodische Schwerpunktsetzungen in der Gestaltung der Verfahren berücksichtigt werden. Zweitens wollten wir unsere Auswahl auf solche Verfahren beschränken, die für Kinder und/oder Jugendliche entwickelt worden bzw. bei diesen anwendbar sind und - unter sonst gleichen Bedingungen - dabei deutschsprachige bevorzugen. Damit sollte eine engere, homogenere Vergleichsgrundlage sowohl für die einzelnen Verfahren untereinander als auch für das von uns vorgestellte geschaffen werden.

Tabelle 1 gibt eine Übersicht über die dabei zustandegekommene Auswahl wieder, auf die wir nun näher eingehen. Unseren Ausführungen möchten wir jedoch noch einige Anmerkungen zu Tabelle 1 voranstellen. In Spalte 1 sind - soweit vorhanden - die eingeführten Abkürzungen eingetragen. In Spalte 2 wurden für das Verfahren PFT die Autoren der deutschsprachigen Adaptation angegeben; der Autor des Originals ist ROSENZWEIG (Näheres im dazugehörigen Textteil). In Spalte 3 wurde letztgenanntes Verfahren als "Persönlichkeits-Entfaltungsverfahren" eingeordnet, weil dies von seiner theoretischen Grundlegung her mehr seinem eigentlichen Charakter entspricht. Daneben weist es fragebogenähnliche Anteile in seiner Gestaltung auf, weswegen es bisweilen auch als "semiprojektiv" bezeichnet wird.

ERHEBUNGSVERFAHREN (Abkürzung)	AUTOR(EN)	TESTMETHODISCHES KONZEPT
HT	BRICKLIN et al. (1962)	PE
FHT	BELSCHNER et al. (1971)	PE
RCT	ELIZUR (1949)	PE
L-TAT	LESSER (1958 a,b)	PE
PFT	DUHM & HANSEN (1957)	PE
PFK	SEITZ & RAUSCHE (1976)	FB
HAPEF-K	WAGNER & BAUMGÄRTEL (1978)	FB
SÜLLWOLD-SKALEN	SÜLLWOLD (1959, 1961)	FB
EAS	PETERMANN & PETERMANN (1980)	FB
HAVEL	WAGNER (1981)	RA
BAV	PETERMANN & PETERMANN (1978)	RA
STEREOSKOP	MOORE (1966)	BTT

PE : Persönlichkeits-Entfaltungsverfahren

FB : Fragebogen

RA : Rating

BTT: Binokulare tachistoskopische Technik

Tab. 1: Übersicht über die zur Erfassung von Aggression und Aggressivität im Kindes- und Jugendalter vorgestellten Erhebungsverfahren.

BRICKLIN, PIOTROWSKI & WAGNER (1962) veröffentlichten ein projektives Verfahren für Kinder, Jugendliche und Erwachsene mit dem Anspruch, damit insbesondere eine Vorhersage offen aggressiven Verhaltens zu ermöglichen, den HAND-TEST (HT).

Das Verfahren besteht aus 10 etwa 8.5 cm x 11 cm großen Tafeln, auf die, bis auf eine, die leer bleibt, je eine Hand in unterschiedlicher Stellung gezeichnet ist. Dem Probanden werden die Tafeln nacheinander vorgelegt mit der jeweiligen Bitte um Angabe, was die Hand seiner Meinung nach gerade tun könnte. Er kann die Tafeln dabei drehen, die Zahl seiner Antworten ist nicht begrenzt.

Das theoretische Grundkonzept des Verfahrens beruht auf der besonderen Bedeutung der Hand für das menschliche Verhalten und Handeln. Aus den Handlungen, die der Proband den Händen auf den Testbildern zuordnet, wird unmittelbar, ohne besondere theoretische oder gar aggressionstheoretische Annahmen auf ihm selbst immanente Handlungstendenzen und mittels dieser auf die Art seines eigenen aggressiven Verhaltens geschlossen.

Aggression wird gescort, wenn Hände als dominierend, verletzend, angreifend oder aktiv einen anderen Organismus oder ein anderes Objekt ergreifend beschrieben werden (ebd., S. 6). Daneben wird die Annahme formuliert, daß direktive Handlungen, die den Händen zugeordnet werden ebensowenig wie die bereits beschriebenen eine Bereitschaft erkennen lassen, sich mit anderen (Mitmenschen) zu

einigen. Eine zweite Gruppe von Kategorien ist für die Einordnung von Handlungen vorgesehen, von denen man annimmt, daß ihnen, unabhängig von ihrem jeweiligen motivationalen Hintergrund, im Gegensatz dazu ein sozial-kooperatives Moment gemeinsam ist. Von der Differenz der Summenwerte beider Kategoriengruppen ausgehend glaubte man nun, mit diesem Index , "acting-out-score" (AOS) genannt, einen besonders engen Zusammenhang zu offen aggressivem Verhalten herstellen zu können. Wir gehen innerhalb des nächstbeschriebenen Verfahrens noch genauer darauf ein.

Psychometrische Gütekriterien zu den Kennwerten des HT finden wir bei BRICKLIN et al. (1962) nicht; der HT war wohl mehr als ein klinisches Erkundungsverfahren gedacht. Im Arbeitskreis um SELG (GÖTZGER 1965; STEINMETZ 1965; vgl. SELG 1974) wurde der HT an Jungen und Mädchen im Alter zwischen 7 und 15 Jahren überprüft. Als Validitätshinweis konnte ein Zusammenhang zwischen einem Aggressionskennwert des HT (nicht näher bezeichnet) und einem Aggressionskennwert aus einem kombinierten Lehrer- und Gleichaltrigen-Rating aufgefunden werden (vgl. SELG 1974, S. 118).

BELSCHNER, LISCHKE & SELG (1971) unternahmen nun auf Grundlage des HT eine Neukonstruktion und legten diese als FOTO-HAND-TEST (FHT) vor. Anlaß dazu waren vor allem zwei kritische Punkte am HT: erstens der "Omnibus-Charakter" des Verfahrens (der HT war nicht nur zur Untersuchung eines Bereichs der Persönlichkeit vorgesehen) und zweitens dessen mangelhafte psychometrische Fundierung (ebd., S. 19)

Man war nun bestrebt, den HT zu einem engdimensionierten, nur für Aggression vorgesehenen projektiven Verfahren umzubauen, das psychometrischen Standards genügte. Da man die Antwortzahl pro Bild auf 1 begrenzte, wurde gleichzeitig die Anzahl der Bilder erhöht, um eine hinreichend breit gestreute Verhaltensstichprobe im Test herstellen zu können. Der FHT besteht nun aus 34 etwas unscharfen Schwarzweißfotos (Format: 9 cm x 12 cm) und nicht mehr aus handgezeichneten Bildtafeln. Eine Leertafel, wie beim HT, gibt es ebenfalls nicht mehr. Darüber hinaus wurden Modifikationen in Vorgabe und Auswertung eingeführt. So darf das Bild etwa nicht mehr gedreht werden.

Das theoretische Grundkonzept des HT behielt man jedoch bei. Vom HT wurde ferner der Hauptindex für offen aggressives Verhalten, der AOS übernommen. Wie bereits zur theoretischen Begründung des AOS bei BRICKLIN et al. (1962), so wird auch hier ganz pauschal angenommen, daß die Wahrscheinlichkeit, sich offen aggressiv zu verhalten, in dem Maße steigt, in dem dominante und aggressive Einstellungen gegenüber solchen, die auf kooperative Aspekte im sozialen Umgang miteinander hinweisen und solchen, die Furcht vor Gegenaggression und Vergeltung beinhalten, die Oberhand gewinnen.

In den AOS gehen die folgenden Antworttendenzen mit dem jeweils angegebenen Vorzeichen ein:
aggressive (agg, +; Antworten, die ein gegen einen Organismus oder Organismus-Surrogat gerichtetes Austeilen schädigender Reize

beinhalten; z.B. "jemanden an den Haaren ziehen"), direktive (dir, +; Antworten, in denen eine Einflußnahme auf den (Handlungs-) Partner erkennbar wird; z.B. "Polizist regelt den Verkehr"), kommunikative (kom, -; Antworten, in denen ein Austausch von Information erkennbar wird, wobei dem Empfänger der Botschaft ein offener Handlungsspielraum erhalten bleibt; z.B. "einen Brief schreiben"), affektive (aff, -; Antworten, die positive Gefühle gegenüber dem Partner erkennen lassen; z.B. "jemanden beruhigen"), Abhängigkeit (dep, -; Antworten, die das Einnehmen einer untergeordneten Position indizieren sowie Bitte um Hilfe und Beistand erkennen lassen; z.B. "Trinkgeld erbitten") und schließlich Furcht beinhaltende (fur, -; Antworten, die Angst und Furcht ausdrücken oder erkennen lassen, daß der Träger der Hand Angriffen oder Bestrafung ausgesetzt ist oder versucht, solche abzuwehren; z.B. "Abwehrhaltung bei Schreck").

Durchführungs- und Auswertungsobjektivität des FHT können als befriedigend, seine Reliabilität dagegen muß als mäßig bezeichnet werden. Zur Dimensionalität der Meßwerte finden wir den Hinweis, daß die mit dem Verfahren erfaßte Form der Aggressivität direkt, zuschlagend und besonders charakteristisch für männliche Jugendliche sei. Zur Validität finden wir u.a. folgende Hinweise: In einer unausgelesenen Stichprobe männlicher Volksschüler korreliert der AOS sehr signifikant mit dem Kennwert aus einem Lehrer-Rating für Aggression, jedoch nicht signifikant mit dem parallel erhobenen Gleichaltrigen-Rating für Aggression als Außenkriterium.

In Übereinstimmung mit dem theoretischen Konzept (mit fortschreitender Sozialisation überwiegen die prosozialen Antriebe) nimmt der AOS mit zunehmendem Alter der Probanden ab und überwiegen (aufgrund geschlechtstypisch unterschiedlich geprägter Sozialisationsbedingungen) die offen-aggressiven Verhaltensanteile beim Mann: LISCHKE (1969; zit. n. BELSCHNER et al. 1971, S. 33) fand bei 10 bis 16 Jahre alten Heimkindern eine signifikant negative Korrelation zwischen den Variablen Alter und AOS; ein AOS-Mittelwertevergleich zwischen männlichen und weiblichen Jugendlichen zeigte eine signifikante Differenz an, der zufolge männliche Jugendliche einen höheren AOS aufweisen als weibliche (vgl. BELSCHNER et al. 1971, S. 34).

BELSCHNER et al. (1971) sind mit dem Vorsatz angetreten, aus dem HT einen richtigen Test zu machen. Wir wollen das Ergebnis (FHT) nun einmal kritisch unter die Lupe nehmen. Die inhaltlich geäußerten Bedenken gelten auch für den HT. Die Kritikpunkte an den formalen Qualitäten des FHT erübrigen sich beim HT, da, wie gesagt, bei diesem ja keine angegeben sind. Summarisch kann man demzufolge zum HT sagen, daß kein Mensch weiß, was er eigentlich wie gut oder wie schlecht erfaßt.

Die Induktionslogik des FHT mutet naiv-direkt an. Was der Proband der Hand zuordnet, soll auch für ihn charakteristisch sein. Überraschenderweise hat einer der drei Autoren des FHT, SELG (1968; zit. n. CHRISTOPH-LEMKE 1972, S. 47), diese Art der Schlußfolgerung

bereits 3 Jahre vor Erscheinen des FHT, allerdings im Zusammenhang mit dem als nächstes Verfahren dargestellten RCT, kritisiert. Wichtiger scheint uns folgender Gesichtspunkt: Die Iteminhalte für Aggression sind nicht aggressionstheoretisch begründet. Man sieht das sehr deutlich, auch in seinen Auswirkungen, an der Miteinbeziehung einzelner Kategorien in den AOS. Gibt man etwa die Antwort "Polizist regelt den Verkehr", steigt der AOS an. Sieht man einmal davon ab, daß direktive Tendenzen eher etwas mit der Ausübung von Macht, einem gänzlich anderen Motivbereich zu tun haben (vgl. HECKHAUSEN 1980), wird dabei noch nicht einmal unterschieden, ob der Proband die o.g. Antwort aus der Perspektive eines Polizisten oder eines Verkehrsteilnehmers gibt, d.h. das Dirigieren als selbst aktiv ausübend oder in seinen Auswirkungen passiv erfahrend erlebt. Dies scheint nun aber doch entscheidend wichtig zu sein, denn je nachdem müßte man die Antwort - der AOS-Logik zufolge - einmal mit einem positiven, das anderemal mit einem negativen Vorzeichen verrechnen. Das heißt also: die theoretisch-spekulative Begründung des AOS bleibt fraglich. Daneben wird auch keine empirische Begründung des AOS (etwa kanonische Variable) gegeben.

Betrachten wir nun einmal näher die formalen Qualitäten der Meßwerte des FHT. Die angegebenen Retestmessungen wurden an einer nur sehr kleinen Stichprobe ($n = 14$) vorgenommen und dazu noch nach einem Zeitintervall, bei dem man kaum abschätzen kann, in welchem Ausmaß es die Enge des Zusammenhangs zwischen Erst- und

Zweitmessung artefiziell in die Höhe treibt (4 Tage). Andererseits sind die zudem unvollständig angegebenen split-half-Koeffizienten bei den normalen Jugendlichen gänzlich unzureichend, die theoretische Begründung ihrer Berechnung zugegebenermaßen fraglich (vgl. BELSCHNER et al. 1971, S. 28 f.). Die Angaben zur Dimensionalität der FHT-Meßwerte (ebd., S. 40 f.) sind unvollständig und lassen kein theoretisch begründetes Strukturkonzept erkennen, dagegen methodische Fehler, was das Einhalten metrischer Voraussetzungen zur Eingabe von Variablen in eine Faktorenanalyse angeht. Aus dem unklaren theoretischen Überbau des FHT schließlich ergeben sich keine Ansatzpunkte zur Ableitung von Beurteilungsgesichtspunkten, nach denen experimentelle "Validierungsbefunde" überhaupt als für oder gegen die Validität des Verfahrens sprechend eingeordnet werden können. Daher sind die angegebenen Befunde nicht sonderlich überzeugend. Es ist nach wie vor unklar, was der FHT erfaßt.

ELIZUR (1949) entwickelte ein System zur Signierung von Angst und Feindseligkeit im Deutungsinhalt von RORSCHACH-Bildtafeln, den projektiven RORSCHACH CONTENT TEST (RCT). Der RCT wurde von STEINMETZ (1965; zit. n. SELG 1974) an 16 Jungen (Durchschnittsalter: 10;9 Jahre) geprüft und wird mit aus diesem Grund hier vorgestellt.

Bei den RORSCHACH-Bildtafeln handelt es sich um die Darstellung von 10 sorgfältig ausgesuchten, zufällig zustandegekommenen symmetrischen Klecksgebilden. Diese sind teils schwarzweiß, teils mehrfarbig.

Aufgrund eigener Untersuchungen war RORSCHACH (1921) zu der Überzeugung gelangt, daß sich aus formalen und inhaltlichen Aspekten der Deutung der Bilder Rückschlüsse auf strukturelle und dynamische Charakteristika der Persönlichkeit des Deutenden ziehen ließen und hatte den Deutungsprozeß selbst in Termini der Wahrnehmung und Auffassung beschrieben (ebd., S. 16 ff.).

Beim RORSCHACH-Test erhält die Versuchsperson die Tafeln nacheinander in die Hand und wird dabei jeweils gefragt: "Was könnte dies sein?" Sie darf die Tafeln dabei drehen und wenden und sollte zu jeder mindestens eine Antwort geben. Nach oben ist die Zahl der Antworten nicht begrenzt.

Das hinter dem RCT stehende theoretische Konzept beruht auf einem tiefenpsychologischen, triebdynamischen Verständnis von Persönlichkeit, in das zudem Anleihen aus der Feldtheorie Kurt LEWIN's eingeflochten sind. Die beiden Persönlichkeitskonstrukta, zu deren Quantifizierung der RCT entwickelt wurde, Angst (anxiety) einerseits und Feindseligkeit (hostility) andererseits, sind als personimmanente Spannungssysteme konzipiert, die - ihrerseits Produkt der Integration vielfältigster Aspekte der Persönlichkeit (Bedürfnisse, Strebungen, Emotionen etc.) und gleichermaßen Ausdruck eines gestörten innerorganismischen Gleichgewichts - steuernd in die Wahrnehmung der Bildtafeln eingreifen und ihren Deutungsinhalt mitbestimmen. Und zwar über die psychischen Mechanismen der Externalisation (eine Person findet außerhalb liegende Gründe für ihre eigene Angst

und Feindseligkeit; z.B. überall Gefahrenquellen sehen und dergl.) und Identifikation (die Person formt Wahrnehmungsinhalte, mit denen sie sich identifiziert; z.B. kämpfende Tiere).

Im Gegensatz zur tradierten RORSCHACH-Test-Auswertung, die mehr Gewicht auf die formalen Aspekte der Deutung legt, nimmt ELIZUR ausschließlich auf deren Inhalt Bezug, den er mit den Kategorien "Angst" (A, anxiety; lt. Definition ein innerer Zustand von Unsicherheit) und "Feindseligkeit" (H, hostility; lt. Definition Gefühle von Ärger und Feindschaft) signiert. Alle übrigen Reaktionen werden in einer neutralen Sammelkategorie ("neutral") aufgefangen. Sind die Deutungsinhalte eindeutig kategorisierbar, werden Großbuchstaben (A, H) vergeben, die doppelt zählen. Sind die Signierungsrichtungen nur abgeschwächt anwendbar, werden Kleinbuchstaben (a, h) benutzt. Für bestimmte Fälle ("double connotation") sind Mischsignierungen vorgesehen (z.B. wird "Polizist" mit ah signiert). Unter den Signierungsbeispielen finden wir bei ELIZUR u.a. folgende: für Angst (A) Wahrnehmungsinhalte wie "Drache", "Monster", "Atombombe"; für Feindseligkeit (H) Wahrnehmungsinhalte wie "Gewehr" oder "zwei Tiere, die miteinander kämpfen". Beispiele für die Vergabe von Kleinbuchstaben sind etwa "Spinne", "Kirche", "Priester" (alle a) oder "Klatschbasen" (h). Die gewichteten Punktwerte der vergebenen Signierungen werden anschließend zu den RCT-Variablen "h" ("total h") für Feindseligkeit und "a" ("total a") für Angst aufaddiert.

Zur Auswertungsobjektivität des RCT gibt ELIZUR Korrelationskoeffizienten zur Übereinstimmung zwischen zwei Auswertern an, die zwischen

.77 und .98 variieren. Reliabilitätsschätzungen (split-half, aufgewertet nach SPEARMAN-BROWN; n = 30) für die RCT-Variablen streuen zwischen .48 (für a) und .75 (für h). ELIZUR meint, klinischen Beobachtungen zufolge sollten die Meßwerte a und h schwach positiv korreliert sein, da Angst und Feindseligkeit oft koinzidierten und einander wechselseitig verstärkten. Die Annahme schwach positiver Korrelation konnte er mit den von ihm vorgelegten Daten bestätigen.

Der RCT stammt aus einer Periode der Testpsychologie, in der man noch glaubte, auf der Grundlage eines einzigen Verfahrens ein vollständiges Bild der Gesamtpersönlichkeit eines Menschen zeichnen zu können. Insofern gilt hier ähnliches, wie bereits in der Kritik zum HT/FHT geäußert, insbesondere hinsichtlich der Induktionslogik des RCT. Sein spekulativ-impressionistisch erstelltes Signierungssystem ist zudem logisch schwer nachvollziehbar, wenn nicht inkonsistent. Die Begründung, "Atombombe" mit "A", "Gewehr" dagegen mit "H" zu signieren ist, daß, aufgrund eines kulturspezifischen Furchtstereotyps, bei ersterer der Furchtaspekt weit überwiegt. ELIZUR übrigens hat diese Unzulänglichkeiten klar gesehen, wie er in einer Fußnote (ELIZUR 1949, S. 258) bemerkt. Unter Zuhilfenahme einer ausreichend großen und repräsentativen Stichprobe sowie nach Durchrechnung einer Itemanalyse wären dann womöglich auch Vorschläge unterblieben, wie "Priester" oder "Kirche" mit "a" zu signieren. Aufgrund der fehlenden Orientierung an psychometrischen Kriterien während der Phase der Entwicklung des Signierungssystems verwundert die Angabe von Reliabilitätsschätzungen für die RCT-Variablen, zumal bei einer Stichprobengröße

von n = 30. Aufgrund der theoretisch als auch empirisch unklaren Situation beim Test-Splitting lassen sich die angegebenen Werte nicht interpretieren. Die Dimensionalität der RCT-Variablen ist nach wie vor unklar, da nicht geprüft. Über eine schwach positive Korrelation läßt sich in dem angeführten Begründungszusammenhang nicht auf Validität schließen. Zudem ist letzterer logisch nicht stichhaltig. Daß die Validität des RCT ungesichert ist, läßt eine Untersuchung von STEINMETZ (1965; nach SELG 1974, S. 106 f.) erkennen: in ihr konnten weder Beziehungen der Meßwerte des RCT zu vergleichbaren aus anderen Aggressionstests noch zu einem Außenkriterium für Aggression (kombiniertes Lehrer- und Gleichaltrigen-Rating) aufgefunden werden.

LESSER (1958 a, b) unternahm den Versuch, in Anwendung und Fortführung der Grundidee des THEMATIC APPERCEPTION TEST (TAT) von MURRAY (1943) einen speziellen Aggressions-TAT (im folgenden "LESSER-TAT" oder kurz "L-TAT" genannt) für Jungen im Alter von etwa 10 bis 13 Jahren zu entwickeln, der valide Messungen für den Bereich "Gleichaltrigen-Aggression" ("peer aggression") erlaubt.

Die Grundidee des TAT läßt sich etwa wie folgt umschreiben: Das Verfahren besteht aus einer Reihe schwarzweißer Bildtafeln und einer leeren weißen. Die meisten davon zeigen eine oder mehrere Personen in Ausschnitten sozialer Situationen. Die Darstellung ihrer Tätigkeiten ist so gehalten, daß sie unterschiedliche Interpretationen zuläßt. Aufgabe des Probanden ist es nun, zu den einzelnen Bildern

möglichst interessante und dramatische Geschichten zu erfinden. MURRAY nahm nun an, daß bei der Interpretation der mehrdeutigen Situation der Proband dazu neigt, Charakteristisches seiner Person ebenso einzubringen wie das darin zu sehen, was objektiv dargestellt ist. So greifen in die Gestaltung der Geschichte die unbewußten Ängste, Wünsche und Konflikte des Probanden ein. Er "projiziert" sie in den dargestellten Handlungsablauf der Geschichte, wo sie, ihre Spuren hinterlassend, des diagnostischen Zugriffs harren. MURRAY formulierte noch weitere Annahmen dazu, auf die jedoch an dieser Stelle nicht weiter eingegangen werden soll (vgl. KORNADT & ZUMKLEY 1982).
Der TAT war als "Breitbandverfahren" zur inhaltsanalytischen Erfassung einer Vielzahl von Motiven (needs) und umweltwirksamen Kräften (presses) gedacht, die in ihrem Zusammenspiel sogenannte "Themata" bilden. Er diente einem mehr klinisch-motivierten Erkundungsinteresse und war psychometrisch nicht fundiert.

LESSER folgte nun einer von AULD, ERON & LAFFAL (1955) vorgezeichneten Forschungsrichtung, die man als Anwendung psychometrischer Methoden auf Analyse und Interpretation projektiven Materials (TAT-Daten) umschreiben kann (LESSER 1958 b). Testpsychologisch läßt sich LESSER´s Entwicklung desweiteren dadurch kennzeichnen, daß er versucht hat, der Vorstellung einer im Grunde konflikthaften Motivation gerecht zu werden. In der Analyse des projektiven Materials unterscheidet er zwischen Annäherungs- ("approach") und Meidenstendenzen ("avoidance tendencies") und nimmt an, daß offen

aggressives Verhalten unmittelbar vom Verhältnis der Ausprägungsgrade dieser beiden Tendenzen abhängt (LESSER 1958 a). Auf diesem Forschungssektor lagen bereits einige Untersuchungen und Erfahrungswerte vor, auf die er sich bezog (PITTLUCK 1950; MUSSEN & NAYLOR 1954; PURCELL 1956; zit. n. LESSER 1958 a).

Im einzelnen besteht der L-TAT aus einem Auswertungsschlüssel für Aggression (FA: fantasy aggression) und Aggressions-Hemmung (AA: aggression anxiety) sowie einem Satz von 10 Bildern. Auf den handgezeichneten Schwarzweißbildern sind jeweils zwei Jungen vergleichbaren Alters in verschiedenen Situationsausschnitten ihres sozialen Alltags (Freizeit, Schule) dargestellt. In die Gestaltung der Bilder sind Ausdruckskomponenten einbezogen, die dem thematischen Bereich "Aggression" zugeordnet werden können (heruntergezogener Mundwinkel, geballte Faust etc.); andere thematische Hinweisreize suchte man zu vermeiden. Der Bildersatz verfügt demzufolge über eine abgestufte Anreizwerteverteilung für Aggression (FA), die zwischen 7 % und 94 % variiert. Dem Probanden werden die Bilder mit der Bitte vorgelegt, zu jedem eine Geschichte zu erfinden, die er nach eigener Phantasie ausgestalten kann. Es gibt keine richtigen und keine falschen Geschichten; er soll einfach nur schreiben, was die Jungen auf den Bildern denken und fühlen und wie die Geschichte schließlich ausgeht.

FA wird gescort, wenn in den Geschichten angegriffen, gekämpft, Unrecht getan, gehaßt wird etc., wobei nur der erste aggressive Akt in jeder Geschichte zählt. Auf diese Weise wird jeder Geschichte ein FA-Wert von 0 oder 1 zugeordnet (FA-Range des L-TAT: 0 bis 10).

Alle Akte von Gegenaggression oder Strafe für Aggression werden als AA gescort. Zusätzlich erfaßt LESSER, ob eine "Instigation" zur Aggression in der Geschichte vorliegt. Dies ist der Fall, wenn ein Junge Bedingungen herstellt, die dazu dienen, beim anderen (Handlungspartner) Aggression zu provozieren. Danach kann Aggressions-Hemmung an unterschiedlichen Stellen des weiteren Interaktionsverlaufs (der beiden Handlungspartner) intervenieren. Die ersten drei der vier diskreten Hemmungskategorien ergeben sich demzufolge durch Differenzierung der Hemmung gemäß deren Auftretenszeitpunktes in einer Instigation von Aggression beinhaltenden TAT-Geschichte (AA prevents, - interrupts, - follows aggression); die vierte nennt er "psychological distance". Sie wird immer dann vergeben, wenn eine Geschichte Hinweise darauf beinhaltet, daß sich der Proband von ihrem Inhalt ("psychologisch") distanziert (z.B. der Proband plaziert sich selbst an den Rand der Geschichte ("story action")). Im Gegensatz zu FA kann für AA pro Bild und Geschichte ein Punktwert > 1 vergeben werden, und zwar insofern, als die gleiche AA-Kategorie mehrmals vergeben wird und/oder mehrere AA-Kategorien signiert werden.

LESSER berechnet nun mit seinem Verfahren zunächst ein Intensitätsmaß für eine generelle Aggressionsbereitschaft (FA) und neben Intensitätsmaßen für die vier unterschiedlichen Arten der Aggressions-Hemmung dann noch eines für einen Gesamt-Hemmungswert ("total anxiety"). Aus diesen Werten (FA, AA_I bis AA_V) leitet er durch Quotientenbildung (FA / AA_i) fünf verschiedene "Konflikt"-Indizes für Aggression ab.

LESSER (1958 a) gibt zur Übereinstimmung zwischen zwei unabhängigen Auswertern Werte zwischen 80 % (psychological distance) und 96 % (FA) an, zur Reliabilität (corrected matched-half) Korrelationskoeffizienten (PEARSON) zwischen .63 (psychological distance) und .83 (FA).

Aufgrund seiner sorgfältigen Entwicklung kann LESSER (1958 b) für den FA-Wert des L-TAT mittels GUTTMAN-Skalierung Eindimensionalität und ordinales Meßniveau nachweisen (Reproduzierbarkeitskoeffizient von .91). Den Versuch, analog dazu auch die AA-Werte zu überprüfen, unternahm LESSER nicht, wies aber auf die Dringlichkeit dieser Weiterentwicklung hin. Man darf daher wohl annehmen, daß LESSER die Skalierbarkeit der Aggressions-Hemmung versucht hat, dabei jedoch nicht zu positiven Resultaten gekommen ist. Die Interkorrelationen der einzelnen Kennwerte sind (leider) nicht aufgeführt.

Zur Konstrukt-Validität des FA-Maßes fand LESSER (1957) einen gleichgerichteten Zusammenhang zwischen diesem und einem Außenkriterium für Aggression (über ein Gleichaltrigen-Rating, eine "Guess who"-Technik erhoben) bei Jungen, deren Mütter Aggression positiv gegenüberstanden ($r = +.43$, sig. 5 %, zw., $n = 23$); bei solchen, deren Mütter Aggression relativ ablehnend gegenüberstanden, drehte sich der Zusammenhang um ($r = -.41$, sig. 10 %, zw., $n = 21$). Die beiden Koeffizienten unterscheiden sich zudem signifikant voneinander (sig. 0.6 %, zw.).

In einer weiteren Untersuchung fand LESSER (1958 a) einen signifikanten Zusammenhang zwischen dem Konflikt-Index FA/AA_I (anxiety prevents) und dem gleichen Außenkriterium von r = +.50 (sig. 1 %, zw., n = 72).

Der L-TAT stellt in mehrfacher Hinsicht eine bemerkenswerte Neuentwicklung dar. Neben der Abkehr vom breit angelegten TAT und der Entwicklung eines engdimensionierten, nur für einen einzigen Motivbereich (Aggression) Gültigkeit beanspruchenden Verfahrens ist als positiv hervorzuheben, daß innerhalb des zugrundegelegten Motivbereichs zusätzlich eine situative Orientierung erkennbar wird, die aus der Menge möglicher Handlungspartner eine ganz bestimmte Untermenge (Gleichaltrige) ausgrenzt. Die damit zum Ausdruck gebrachte Einschränkung des Gültigkeitsbereichs der Messung erschwert andererseits natürlich die Ableitung sog. "extensiver" Meßgrößen, mit denen angezeigt werden soll, in wievielen unterschiedlichen Situationsbereichen das Motiv überhaupt anspricht. Der L-TAT erlaubt daher im Grunde nur die Ermittlung "intensiver", für den o.g. Situationsbereich gültiger Meßgrößen. Weiter ist positiv anzumerken, daß der L-TAT für eine engumschriebene Population entwickelt wurde, die gleichermaßen geschlechts- und altersbezogen eingegrenzt ist. Die empirische TAT- und Aggressionsforschung hat seitdem Ergebnisse gezeitigt, die diese Einschränkungen und Besonderheiten als sinnvoll erscheinen lassen (vgl. KORNADT & ZUMKLEY 1982; PARKE & SLABY 1983).

Eine Schwachstelle des L-TAT ist zweifelsohne, daß die im

Annäherungs-Teil des Aggressionsmotivs (FA) geglückte Erfüllung metrischer Erfordernisse im Hemmungs-Teil desselben (AA) noch nicht nachvollzogen werden konnte. Es fehlt zudem die Angabe der Interkorrelationen der Kennwerte des L-TAT in diesem Zusammenhang. Die LESSER'schen "Konflikt"-Maße darf man jedoch nicht als Indikatoren für den Grad der Konflikthaftigkeit der Aggressionmotivation deuten. Durch einfache Quotientenbildung, wie LESSER sie vornimmt, werden nämlich zwei theoretisch wichtige Parameter motivationaler Konflikte, die Intensitätsdifferenz der konfligierenden Motivierungen als auch die Stärke der Gesamtmotivation, die sie bilden, nicht berücksichtigt geschweige denn sinnvoll verrechnet (vgl. KASTNER 1977, S. 226; WESTHOFF 1977, S. 161 ff.).

DUHM & HANSEN (1957) legten die deutschsprachige Adaptation der Kinderform eines semiprojektiven Verfahrens vor, der als "Aggressions"-Test weltweit bekannt wurde, den ROSENZWEIG PICTURE FRUSTRATION TEST (PFT). Es handelt sich dabei um eine im Test simulierte Reaktionsprüfung auf belastende (frustrierende) Situationen des Alltags.

Der Proband findet im PFT 24 schwarzweiße Strichzeichnungen (Format etwa 8 cm x 10 cm) vor, in denen Ausschnitte aus verschiedenen sozialen Situationen des Alltags dargestellt sind. Die in der linken Bildhälfte dargestellte Person (mal Erwachsener, mal gleichaltrige oder ältere Kinder) teilt dabei der rechts davon dargestellten (Kind; abwechselnd Junge oder Mädchen) in einer Sprechfahne

etwas mit, das, so die Annahme, für diese eine milde Frustration beinhaltet. Die Sprechfahne der angesprochenen Person ist leer. Die Gesichter der Personen sind unstrukturiert gehalten, Ausdruckskomponenten wurden vermieden. Durch die kurzen Sätze wird jedoch der Aufforderungsgehalt der Bilder potenziert und den jeweils dargestellten Thematiken ein hoher Strukturierungsgrad verliehen.

Die Aufgabe des Probanden besteht nun darin, die Bilder der Reihe nach zu betrachten und den eingetragenen Text zu lesen. Darauf soll er in die leere Sprechfahne stets die erste ihm einfallende Antwort eintragen, die seiner Meinung nach die angesprochene Person geben könnte. Bei jüngeren Kindern protokolliert der Testleiter die Antworten mit.

Der PFT ist unmittelbar aus der Frustrationstheorie von ROSENZWEIG (1934, 1938, 1944; zit. n. HILTMANN & LUHR 1982) abgeleitet und war zu deren empirischer Prüfung gedacht. Er war dagegen nicht zur Erfassung gar funktionell-begründeter Konstrukta und deren Quantifizierung vorgesehen, wurde jedoch in der Folge weitgehend so aufgefaßt und (meist) unkritisch übernommen (z.B. LESSER 1959 a). Eine Annahme des PFT besteht nun darin, daß der Proband jede Testsituation (als für die angesprochene Person) frustrierend auffaßt, eine weitere darin, daß er sich anschließend mit der frustrierten Person identifiziert. Unter "Frustration" versteht ROSENZWEIG ein psychodynamisches Phänomen, das auftritt, wenn eine Person mit Hindernissen konfrontiert oder die Befriedigung ihrer vitalen Bedürfnisse

gehemmt wird. Ferner wird angenommen, daß es sich bei den Antworten des Probanden, die er der frustrierten Person zuschreibt, um Projektionen eigener Einstellungen, Haltungen und Handlungstendenzen handelt, die aggressiver Natur sind. Ihre Auswertung liefert Aufschluß über Art und Richtung der durch die Frustration bedingten Aggression, die der Proband selbst in vergleichbaren Situationen seines sozialen Alltags zeigt. Dabei werden drei Richtungen unterschieden. Die Aggression kann "extrapunitiv" sein (sich gegen die Umgebung richten), "intropunitiv" sein (sich gegen das eigene Ich richten) und "impunitiv" sein (umgangen werden). Der daran geknüpfte Begriff der "Frustrationstoleranz" bezeichnet die Fähigkeit einer Person, auf Frustrationen in einer sozial angemessenen Weise zu reagieren; sie hängt ab vom Grad der intellektuellen Ausstattung und (sozialen) Reife. In der Antwortreaktion kann weiter das Hindernis dominieren ("obstacle-dominance"), der Bezug auf das Ich überwiegen ("ego defense") oder die Lösung der problematischen Situation betont und/oder angestrebt werden ("need-persistence"). Da jede Testantwort hinsichtlich Typ und Richtung ausgewertet wird, ergibt sich daraus ein System von 9 Hauptsignierungsfaktoren, deren Anwendung anhand einiger ausgewählter Beispiele kurz verdeutlicht werden soll.

Die erste Bildsituation etwa zeigt eine Frau, die einem Mädchen, das, auf einem Stuhl stehend, in einem Schrank offensichtlich etwas sucht, mitteilt: "Ich habe das letzte Stück deinem Bruder gegeben." Extrapunitive Antworten darauf sind dann etwa "Oh" (mit

dem Aspekt obstacle-dominance), "Warum hast du es mir nicht gegeben?" (mit dem Aspekt ego-defense) und "Hast dur noch eins?" (mit dem Aspekt need-persistence). Eine impunitive, ego-defensive Antwort darauf wäre etwa "Ja, es ist gut."

Zur Objektivität des PFT teilen DUHM & HANSEN (1957) nach einer Auswertung von 100 Testheften in 89 % der Fälle Übereinstimmung mit, nachdem die Autorinnen die ersten 50 Hefte gemeinsam ausgewertet und besprochen hatten. Zu weiteren psychometrischen Qualitäten der Meßwerte des PFT finden wir bei DUHM & HANSEN keine Angaben. Einen recht umfangreichen Überblick über Studien zur Validität des PFT als auch über formale Qualitäten seiner Kennwerte gibt allerdings RAUCHFLEISCH (1979).

Den PFT als "Test" einer Kritik zu unterziehen fällt schwer; und zwar zunächst aus dem Grund, weil er von ROSENZWEIG selbst und seinen Mitarbeitern garnicht als Test, sondern als eine kontrollierte, projektive Methode ("Study") gedacht war, als ein Forschungsinstrument, um ROSENZWEIG´s Frustrationstheorie empirisch zu prüfen. ROSENZWEIG´s Frustrationstheorie war jedoch ein mehr spekulatives, an der psychoanalytischen Lehre des Lust-/Realitätsprinzips orientiertes Konzept und hat zu einer differenzierten Analyse des Frustrations-Reaktions-Prozesses nichts Wesentliches beizutragen vermocht (vgl. KORNADT 1981 a). Im PFT wird z.B. angenommen, daß der Proband jede der angebotenen Situationen als frustrierend auffaßt und erlebt. Damit hängt seine (theoretische) Validität unabdingbar

zusammen. Für die Erwachsenenform des PFT konnten WESTERMANN & HAGER (1983) diese Annahme widerlegen. Für die Kinderform des PFT wird sie nach wie vor - empirisch ungeprüft - aufrecht erhalten. Eine weitere Annahme im PFT ist, alle Reaktionen auf Frustration seien als aggressiv charakterisierbar. Die theoretische Entwicklung der Frustrations- und Aggressionsforschung ist hier jedoch einen anderen Weg gegangen. Die durch die F-A-Theorie ursprünglich formulierte Annahme, auf jede Frustration folge eine Aggression, wurde ja schon kurz darauf durch MILLER aufgehoben. Die aggressionsthematische Spezifität der im PFT signierten Reaktionen ist also nicht nur empirisch nicht belegt, sondern steht auch mit speziellen Annahmen der F-A-Theorie nicht in Einklang. Das läßt insgesamt erhebliche Zweifel daran aufkommen, ob das gesamte Verfahren überhaupt zur Quantifizierung bestimmter Aspekte der Aggressivität geeignet ist, zumindest in der Form, in der es vorliegt. Das soll nicht heißen, daß in den Testreaktionen nicht auch Beispiele zu finden wären, die ganz unmißverständlich feindselige oder auch instrumentelle Aggression zum Ausdruck bringen.

Die in sich zum Teil widersprüchlichen und in einzelnen Details auch gänzlich unzureichenden Angaben zur Reliabilität und Validität des PFT stehen bei RAUCHFLEISCH (1979) vergleichsweise konzeptionslos nebeneinander, sodaß wir an dieser Stelle nicht näher darauf eingehen wollen. Mit einigen ungeklärten Problemen setzen wir uns im empirischen Teil dieser Arbeit noch einmal auseinander. Infolgedessen ist die Validität des PFT bis heute sehr umstritten.

BRICKENKAMP (1983) etwa hält sie für erwiesen; gegenteiliger Ansicht sind SELG (1974), SCHMIDTCHEN (1975), HILTMANN & LUHR (1982) oder STEINHAUSEN (1979), der resümierend zu der Auffassung gelangt, " ... den Test in der diagnostischen Praxis und Forschung nicht mehr einzusetzen" (ebd., S. 255).

SEITZ & RAUSCHE (1976) brachten ein multidimensionales Fragebogenverfahren heraus, den "PERSÖNLICHKEITSFRAGEBOGEN FÜR KINDER 9 - 14" (PFK). Es handelt sich beim PFK um einen heterogenen, psychometrisch fundierten Persönlichkeitstest, der zu einer relativ umfassenden Bestandsaufnahme der wesentlichsten Persönlichkeitsmerkmale von Kindern entwickelt wurde.

Der Proband findet im PFK meist in der Ich-Form formulierte Aussagen in der Gestalt kleiner Sätze vor, zu denen er um Stellungnahme gebeten wird. Diese erfolgt durch Ankreuzen in der Antwortspalte "Stimmt" oder in der Antwortspalte "Stimmt nicht". Damit soll er angeben, ob der Inhalt der Aussage für ihn persönlich zutrifft.

Das Testsystem wurde vor dem theoretischen Hintergrund eines multifaktoriellen trait-Konzepts der Persönlichkeit erstellt. Zur Legitimation wird auf diverse, faktorenanalytisch begründete Persönlichkeitstheorien (EYSENCK, CATTELL, GUILFORD etc.; vgl. SEITZ & RAUSCHE 1976, S. 6) verwiesen. Daneben werden aus der Literatur bekannte empirische Untersuchungen zu den in Frage stehenden Persönlichkeitsmerkmalen angeführt und im Zusammenhang mit den

Subskaleninterkorrelationen des PFK diskutiert. Innerhalb der Persönlichkeit differenziert der PFK drei verschiedene Äußerungsformen des Verhaltens, und zwar Verhaltensstile (VS), Motive (MO) und Selbstbild (SB).

Die o.g. Ebenen des Verhaltens wurden der Itementwicklung zugrunde gelegt und führten zur Erstellung von Subskalen für 4 Verhaltensstil-, 6 Motiv- und 5 Selbstbild-Dimensionen. Die Kennwerte auf diesen 15 Primärdimensionen nach den Schwerpunkten ihrer Ladungsvektoren zu Indizes für 5 weitere Sekundärdimensionen zusammengefaßt werden. Anhand ihrer Iteminhalte und Dimensionalität weisen insgesamt 3 PFK-Subskalen für kindliche Aggression unterschiedliche Plausibilität auf, und zwar in abgestufter Reihenfolge MO 1 ("aggressives Bedürfnis nach Ich-Durchsetzung"), MO 6 ("Maskulinität der Einstellung") und VS 2 ("Fehlende Willenskontrolle")

MO 1 wird definiert als "nicht nachgeben, ... , nörgeln, ... , Schadenfreude und Zerstörungswut zeigen." Das erste Item dieser Skala lautet: "Manchmal gebe ich einfach nicht nach, bis ich meinen Willen durchgesetzt habe." Ein aggressionsnahes lautet: "Wenn ich richtig wütend werde, mache ich ab und zu etwas kaputt." Interpretiert wird die der Skala MO 1 zugrundeliegende Testdimension als "aggressiv gefärbte Ich-Durchsetzung unter dem Aspekt fehlender Selbstbeherrschung, weniger unter dem der Feindseligkeit."

MO 6 wird definiert als "Bevorzugung aktiver, aggressiver, realistischer anstelle von feinfühligeren, ästhetischeren, ideelleren Erlebnis- und Selbstbild-Inhalten." Das erste Item dieser Skala lautet: "Ich wäre (a) lieber Lehrer, (b) lieber Anführer einer Armee von Soldaten." Nach diesem Muster sind allerdings nur einige wenige Items formuliert. Der Proband muß hier entscheiden, welche der beiden angebotenen Alternativen für ihn zutrifft. Bei dem o.g. Beispiel indiziert (b) das in Frage stehende Konstrukt. Ein weiteres, mehr auf eine Bereitschaft zu instrumenteller Aggression verweisendes, lautet: "Bevor ich mich einschüchtern lasse, haue ich lieber zu." Interpretiert wird die der Skala MO 6 zugrundeliegende Testdimension im Sinne eines geschlechtsspezifisch anerzogenen bzw. zusätzlich konstitutionell beeinflußten Einstellungs- und Problemlösungs-Komplexes.

VS 2 wird definiert als "fehlende Beherrschung und Selbstkontrolle, Über-Ich-Schwäche, d.h. Schwierigkeiten haben, eigene Bedürfnisse zu unterdrücken, leicht soziale Spielregeln übertreten (Schummeln, Unehrlichkeit)" etc. Das erste Item dieser Skala lautet: "Viele Dinge, die von mir verlangt werden, tue ich ungern." Ein aggressionsnah formuliertes Item lautet: "Ich werde schnell wütend." Interpretiert wird die der Skala VS 2 zugrundeliegende Testdimension als "Bestandteil eines Selbstkonzepts, das aus der kognitiven Auseinandersetzung mit den durch Erziehung und Kultur angebotenen Normen, Überzeugungen und Zielinhalten resultiert." Und weiter: "Aggressive, ich-dominante Bedürfnisse (MO 1) setzen sich durch, sofern

Kontrollmechanismen nicht entwickelt sind, erscheinen möglicherweise auch als Reaktion auf die durch das Erleben fehlender oder unzureichender Identifikationsvorbilder gegebene Frustration."

Der PFK kann als durchführungs- und auswertungsobjektiv bezeichnet werden. Zur Reliabilität (KR 20) sind folgende Werte angegeben: MO 1 (.82), MO 6 (.66), VS 2 (.81). Zur Dimensionalität der in Betracht stehenden Skalen finden wir u.a. folgende Information: MO 1, MO 6 und VS 2 laden den testspezifischen Faktor II. Ordnung "F-II-01" ("derb-draufgängerische Ich-Durchsetzung").

Die Subskalen des PFK mögen für einige Aspekte impulsiv- (MO 1, VS 2) oder auch instrumentell-aggressiven Verhaltens (MO 6) Relevanz besitzen, jedoch kaum für feindselig-motivierte Aggression. Das wird auch im Manual unumwunden zugestanden. Es lag nicht in der Absicht der Autoren, letztere auf evtl. unterschiedlichen Ebenen zu erfassen. Die Itembesetzungen der einzelnen Skalen sind ihren Definitionen entsprechend. Die Interpretation dessen nun, was die einzelnen Skalen tatsächlich erfassen, stützt sich nicht auf empirische Befunde, die mit den Skalen selbst erhoben wurden; man bezieht sich statt dessen mehr auf eine intuitiv-spekulative Interpretation der Iteminhalte und Subskaleninterkorrelationen und sucht davon ausgehend Zusammenhänge zu aus der Literatur bekannten empirischen Befunden herzustellen. Da der PFK insgesamt nicht logisch aus einem theoretisch geschlossenen Persönlichkeitssystem abgeleitet und darin validiert wurde, bleibt die Interpretation seiner Meßwerte insgesamt problematisch und damit auch die o.g. Dimensionierungsangaben (interne Validität).

WAGNER & BAUMGÄRTEL (1978) legten ebenfalls einen Fragebogen vor, den "HAMBURGER PERSÖNLICHKEITSFRAGEBOGEN FÜR KINDER" (HAPEF-K). Das Verfahren ist für Jungen und Mädchen im Alter zwischen 9;0 und 13;11 Jahren gedacht mit dem erklärten Ziel, die bei dieser Klientel für die sozialpädagogische und psychotherapeutische Praxis bedeutsamsten Persönlichkeitsdimensionen zu erfassen.

Zur theoretischen Grundlage des Verfahrens ist implizit erkennbar, daß man - wie beim PFK - auf ein faktorenanalytisch begründetes, multidimensionales Konzept von Persönlichkeitszügen (traits) Bezug nimmt. Bei der Entwicklung der Items orientierte man sich an vorliegenden in- und ausländischen Persönlichkeitsfragebögen sowohl für Kinder als auch für Erwachsene. In den als Feststellungen formulierten Items ist ebenfalls eine enge Orientierung am offenen Verhalten erkennbar.

Unter den insgesamt 6 Subskalen, mit denen u.a. Extraversion, Neurotizismus und Reaktion auf Mißerfolg gemessen werden, ist auch eine für Aggressivität vorgesehen. Mit Aggression " ... ist vor allen Dingen aggressives und auch dominantes Verhalten beschrieben, wobei besonders die Aggression im Kontakt mit Erwachsenen zum Ausdruck kommt." (ebd., S. 18)

Die Auswahl der Feststellungen für die 24-Item-Skala "Aggression" (AG) ist nach Zielobjekten aggressiver Handlungen logisch untergliedert. Neben Aggression gegen Erwachsene finden wir Aggression

gegen Gleichaltrige und nichtbelebte Objekte. Daneben werden Aggressionsphantasien und Items, in denen eine leichte Erregbarkeit zum Ausdruck kommt, als weitere Gliederungsaspekte aufgeführt. Hinsichtlich der Manifestation im offenen Verhalten lassen sich Erscheinungsformen wie Vergeltungsaggression, impulsive Aggression, reaktive, instrumentelle und feindselige Aggression unterscheiden. Daneben gibt es Feststellungen, die, wie schon einige der Subskala MO 6 des PFK, einem attitüdentheoretischen Konzept (HEINRICH 1961, S. 34 ff.; SÜLLWOLD 1959, 1961) entstammen könnten ("Ich habe Lust, auf Großwildjagd zu gehen.") wie auch solche, deren aggressionsthematischer Bezug unklar bleibt ("Ich versuche dem Lehrer Fragen zu stellen, die er nicht beantworten kann.").

Das erste Item der Skala lautet: "Ich treibe mit meinen Freunden gern Unfug." Der Proband nimmt dazu auf einer 6-stufigen Skala Stellung, deren Extrempositionen durch "immer oder fast immer" und "nie oder fast nie" markiert werden. Wie schon beim PFK, so wird auch hier durch Addition der Zustimmungsreaktionen (hier unter Berücksichtigung deren Gewichtung) der Skalenrohwert ermittelt.

Der HAPEF-K kann als durchführungs- und auswertungsobjektiv bezeichnet werden. Die Reliabilität der Skala AG ist gut (interne Konsistenz nach HOYT = .89 bzw. .90 in einer weiteren Untersuchung; die Stabilität beträgt nach 23 Monaten .50). Mit den übrigen Skalen des Fragebogens weist sie maximal 25 % gemeinsame Varianz auf und markiert damit eine eigenständige Dimension. Zur Validität der

Skala AG finden wir u.a. die Hinweise, daß Jungen sich aggressiver als Mädchen darstellen, eine signifikante Korrelation von .19 zur Aggressionsskala der HAVEL (die nachfolgend noch beschrieben wird) sowie eine signifikante Korrelation von .29 zu einem Außenkriterium (Lehrerurteil über die Aggression).

Die Kritik an der Skala AG des HAPEF-K setzt dort ein, wo man versucht hat, zu umschreiben, was sie erfassen soll. Aggression mit aggressivem Verhalten zu erklären ist zirkulär, es mit dominantem Verhalten zu erklären, schwer nachvollziehbar. Danach weiß man nun eigentlich garnicht, was sie überhaupt erfassen soll. Die Itemauswahl trägt dem Rechnung. Durch die Konfusion unterschiedlichster Zielobjekte und Handlungstypen von Aggression erzwingt man praktisch eine oberflächlich-pauschale Interpretation. Aus den psychometrischen Daten der Skala geht zwar hervor, daß sie etwas meßgenau erfaßt. Aus den Angaben zur Validität läßt sich jedoch kaum begründen, daß es sich dabei um eine habituell-instrumentelle oder habituell-feindselige Aggressivität handelt. Die Korrelation zur HAVEL-Skala ist theoretisch unergiebig, da sie von einem ähnlichen Aggressionsverständnis ausgeht und die o.g. Kritikpunkte auf sie ebenso zutreffen. Ähnlich dürfte auch die Beziehung zum Lehrer-Rating zustandegekommen sein, das mit dem unveröffentlichten Fragebogen von BOTT (1971; vgl. WAGNER & BAUMGÄRTEL 1978) erhoben wurde. Bedeutsamer im Hinblick auf eine inhaltliche Aufklärung scheinen u.E. andere Validierungsdaten, die WAGNER & BAUMGÄRTEL mitteilen. So korreliert die Skala AG mit der Skala

"Anstrengungsvermeidung" des AVT (ROLLETT& BARTRAM 1977) in einer Größenordnung von .47 und mit der Skala "Schulunlust" des AFS (WIECZERKOWSKI et al. 1975) in einer ebenso hohen von .46 (vgl. WAGNER & BAUMGÄRTEL 1978, S. 37). "Sinnvollerweise" lautet nun Item Nr. 29 der Skala AG: "Wenn ich im Sommer unbedingt zum Baden gehen möchte, verschiebe ich die Hausaufgaben auf später." Es ist daher nicht auszuschließen, daß die Skala AG mehr Aspekte eines Verhaltenssyndroms erfaßt, das der Mann auf der Straße als "aggressiv" zu etikettieren geneigt ist, als daß es per se etwas damit zu tun hat. Neben der theoretischen Ableitung und Begründung der Skala fehlt auch (vorerst) deren empirische Verankerung an einem validen Kriterium für Aggression.

SÜLLWOLD (1959, 1961) hat - unter einer mehr sozialpsychologischen Perspektive - ein Verfahren zur Erfassung von Einstellung und Bereitschaft zur Aggression bei Jugendlichen beiderlei Geschlechts entwickelt. Das Verfahren liegt in zwei Parallelformen vor und verlangt Stellungnahmen zu Aussagen (Feststellungen) in der Form von vollständigen Sätzen ("Sentenzen-Skala"), Filmtiteln ("Filmtitel-Skala") und Filmbildern, die charakteristische Szenen aus bestimmten Filmen darstellen und zur Werbung für den jeweiligen Film Verwendung gefunden hatten ("Filmbilder-Skala").

Theoretische Grundlage der Skalen ist ein attitüdentheoretischer Bezugsrahmen zur Erklärung von Aggression (vgl. HEINRICH 1961, S. 34 ff.). Die Attitüde "Aggressivität" wird nach diesem

Deutungskonzept als eine mehr oder weniger bewußte, generelle Einstellung gegenüber Aggression betrachtet, als eine konstante Verhaltensbereitschaft, die sich aus individuellen Anlagen, Erziehung und persönlichen Erfahrungen herausgebildet hat. Im Zusammenspiel mit anderen Einstellungen und ggfls. durch diese auch in ihrer Handlungswirksamkeit überlagert, greift sie, variiert durch persönliche Begleitumstände und situative Bedingungen, in die Aktivierung aggressiver Verhaltensmuster ein oder hat ein Aggression billigendes Verhalten zur Folge.

Ein Item der Sentenzen-Skala lautet: "Das Leben ist rauh; darum laß dir nichts gefallen!" Der Proband gibt dazu auf einer 6-stufigen Reaktionsskala mit Ausprägungsgraden von "stark dafür" bis "stark dagegen" seine individuelle Stellungnahme ab. Hier wird bei starker Zustimmung auf hoch ausgeprägte Aggressivität geschlossen. Bei "aggressivitäts-negativen" Items wie "Wenn auch das Schlimmste vorgekommen ist, so mußt du doch verzeihen." wird dagegen starke Ablehnung als Indikator für hohe Aggressivität betrachtet.

Auf der Filmtitel-Skala gibt der Proband an, wie gern er z.B. in einen Film mit dem Titel "Der Turm der Vernichtung" gehen würde, und zwar auf einer ebenfalls 6-stufigen Reaktionsskala, die Ausprägungsgrade von "äußerst gern" bis "äußerst ungern" festhält. Bei diesem Item würde "äußerst gern" hohe Aggressivität indizieren. Bei dem Filmtitel "Hab mich lieb" dagegen würde "äußerst

ungern" auf hohe Aggressivität schließen lassen.

Auf der Filmbilder-Skala gibt der Proband auf einer gleichgestalteten Reaktionsskala an, wie gern er, nachdem er das jeweilige Bild gesehen hat, in den betreffenden Film gehen würde. Auch hier wird - bei friedfertigen Bildern - die Scorung für Aggressivität invertiert, d.h. aus einer stark ablehnenden Reaktion auf hoch ausgeprägte Aggressivität geschlossen. Durch einfache Addition über alle Items wird so pro Skala ein gewichteter Rohwert ermittelt.

Die SÜLLWOLD-Skalen können, wie die anderer Fragebogenverfahren auch, als durchführungs- und auswertungsobjektiv betrachtet werden. Durch die Formulierung inhaltlich entgegengerichteter Items hoffte man zusätzlich, einer möglichen Bildung stereotyper Antwortmuster (Response-Set) entgegensteuern zu können. Zur Dimensionalität der Meßwerte finden wir folgende Angaben: Nach den Ergebnissen von Skalogramm-Analysen sind die Meßwerte der einzelnen Skalen näherungsweise eindimensional (Reproduzierbarkeitskoeffizienten um .80, sog. "Quasi-Skalen"). Die Korrelation zwischen den Film-Skalen beträgt .83; nach Minderungskorrektur wird sie mit .98 angegeben. Die Sentenzen-Skala korreliert mit der Filmtitel-Skala schwach (.31), ebenso mit der Filmbilder-Skala (.21). Zur (Paralleltest-) Reliabilität sind für die einzelnen Skalen Korrelationskoeffizienten zwischen .80 und .87 angegeben. Als Hinweise auf die Validität der Skalen werden verstanden: Männliche Jugendliche

erzielen auf den Skalen höhere Werte als weibliche, jüngere höhere als ältere. Es besteht allerdings ein tendenzieller positiver Zusammenhang zwischen der Häufigkeit des Kinobesuchs (keine Angabe, welche Filme) und der durch die Skalen erfaßten Einstellung. Nach Extremgruppenbildung ("Aggressive" vs. "Verträgliche" nach Gleichaltrigen-Rating) wurden auf den Skalen signifikante Mittelwerteunterschiede in der erwarteten Richtung gefunden. Aggressive hatten höhere Werte als Verträgliche. Man versuchte auch, Validität durch Korrelation der Skalenwerte mit einer nach eigenen Gesichtspunkten errechneten PFT-Extrapunitivität nachzuweisen, konnte jedoch aus technischen Gründen den Zusammenhang zwischen letzterer und der Filmbilder-Skala nicht ermitteln. Zur Filmtitel-Skala zeigte sich dabei eine Korrelation von .10 (n.s.), zur Sentenzen-Skala eine signifikante Korrelation von .29. Aufgrund der empirischen Befunde und in Übereinstimmung mit SÜLLWOLD ist die Annahme wahrscheinlich, daß die beiden Film-Skalen weitgehend Identisches erfassen, die Sentenzen-Skala dagegen etwas anderes.

Nach der Induktionslogik der SÜLLWOLDschen Film-Skalen verweist auf Aggressivität, wenn ein Proband den Film "Der Turm der Vernichtung" oder ähnliche Titel ("Im Zeichen des Zorro" etc.) einer Liebesschnulze vorzieht. In diesem Bevorzugungsverhalten finden u.E. eher geschlechterstereotyp geprägte Interessen ihren Ausdruck, die durch häufigen Kinobesuch noch verstärkt werden können. Es ist kein Wunder, daß auf diese Weise männliche Probanden höhere Werte erzielen als weibliche und jüngere höhere als ältere. Wie steht es

jedoch nach der empirischen Befundlage um die Möglichkeiten, aus den angegebenen geschlechts- und alterstypischen Differenzen in den zur Diskussion stehenden Meßwerten auf deren Validität zu schließen? Eine derartige Beweisführung zur Konstruktvalidität wurde ja schon beim FHT und zum Teil auch bei der Skala AG des HAPEF-K aufgezeigt und wird es, wie wir noch sehen werden, bei MOORE (1966), dort jedoch in umgekehrter Richtung hinsichtlich der Altersdifferenzen. Da all die o.g. Meßwerte nicht weiter differenzieren, läuft die Beweisführung etwa wie folgt: (a) Angehörige des männlichen Geschlechts zeigen mehr Aggression als Angehörige des weiblichen. (b) Mit zunehmendem Alter zeigen die Menschen weniger Aggression. Lassen sich diese Unterschiede in entsprechenden Meßwertereihen nachweisen, gelten sie als Beleg für deren (Konstrukt-) Validität.

ad (a): Nach BARON (1977, S. 218 ff.) ist diese Hypothese empirisch gut fundiert, sodaß man sie zu Recht als Nachweis für Konstruktvalidität eines Verfahrens heranziehen kann. Ihre Gültigkeit scheint dabei allerdings an einen kritischen Punkt geknüpft zu sein: Bei Vorliegen einer starken Provokation treten die bestehenden Unterschiede in den Hintergrund. Da diese Situationscharakteristik jedoch auf die Applikationsbedingungen des Verfahrens nicht zutrifft, bleibt die Triftigkeit der oben genannten Schlußfolgerung bei der Führung des Validitätsnachweises davon unberührt. BARON spricht in diesem Zusammenhang von einem niedrigeren "Siedepunkt" für Aggression beim männlichen Geschlecht, legt im weiteren jedoch auch ebenso plausibel dar,

daß man hinsichtlich bestehender Geschlechtsdifferenzen der Aggression im Zusammenhang mit einem Wandel tradierter Geschlechtsrollenvorstellungen - insbesondere zum Verhalten des weiblichen Geschlechts in der Öffentlichkeit als auch im Zusammenhang mit einem damit einhergehenden, veränderten modalen Persönlichkeitsbild der Frau (u.a. weniger Passivität, dagegen mehr Selbstvertrauen etc.) - in Zukunft durchaus damit rechnen kann, daß diese weitgehend zurücktreten, wenn nicht in Gänze schwinden.

ad (b): Hier sprechen die empirischen Befunde gegen die pauschal formulierte Annahme, jüngere Menschen verhielten sich generell aggressiver als ältere bzw. mit zunehmendem Alter nehme die Aggressivität überhaupt ab. Sie stützen jedoch die Auffassung, daß im Verlauf der individuellen Entwicklung sowohl ein Wandel in den Auslösebedingungen für Aggression als auch in deren phänomenalen Erscheinungsbild stattfindet (vgl. z.B. FESHBACH 1970, S. 184 ff.). HARTUP (1974; zit. n. PARKE & SLABY 1983, S. 569) z.B. fand, daß jüngere Kinder gegenüber älteren mehr instrumentelle Aggression zeigten, letztere jedoch gegenüber ersteren mehr feindselige und gegen Personen gerichtete Aggression. Im übrigen laufen die in dem Übersichtsreferat von OLWEUS (1979) zusammengetragenen Befunde zur Stabilität der Aggression der Vorstellung einer durchgehenden Abnahme mit dem Alter zuwider.

Und weiter: Was kümmert es SÜLLWOLD, die beiden Film-Skalen mit .83 und nach Minderungskorrektur mit .98 korreliert zu sehen? Wäre es da aus testökonomischen Gründen nicht besser, gleich auf eine zu verzichten? Und: Wenn die Sentenzen-Skala mit den beiden Film-Skalen

kaum korreliert ist, was erfaßt sie dann? Aus der angegebenen PFT-Korrelation läßt sich nichts entnehmen. Gofls. dann etwas, wie auch bei den übrigen Skalen, aus der signifikanten Mittelwertedifferenz nach Extremgruppenbildung. Aus der von SÜLLWOLD mitgeteilten Beschreibung der Rating-Kategorien (für "Aggressive" sollte gelten: "angriffslustig, streitlustig, draufgängerisch bis gewalttätig"; für "Verträgliche": "zurückhaltend, versöhnlich, friedfertig bis nachgiebig") läßt sich ebensogut ableiten, daß damit Personen selegiert wurden, bei denen "Ich-Durchsetzung" (vgl. Skala MO 1 des PFK) als Motivation im Vordergrund stand. SEITZ & RAUSCHE (1976, S. 25) glauben von diesem Persönlichkeitszug, daß eine negative Beziehung zu Leistungswerten, die aus Beurteilungen (z.B. durch Lehrer) resultieren, denkbar ist, und zwar übrigens mit der gleichen Begründung, mit der SÜLLWOLD (1959, S. 186) die "... schwache Tendenz zur negativen Korrelation" der Skalen zum intellektuellen Niveau erklärt, "wenn wir das intellektuelle Niveau des Probanden nach bestimmten Schulnoten definierten." Insofern ist hier eine äußerst zurückhaltende Interpretation angebracht und mithin insgesamt unklar, was die SÜLLWOLD-Skalen eigentlich erfassen.

PETERMANN & PETERMANN (1980) legten einen engdimensionierten Fragebogen vor, den "ERFASSUNGSBOGEN FÜR AGGRESSIVES VERHALTEN IN KONKRETEN SITUATIONEN" (EAS). Der EAS ist für Jungen (EAS-J) und Mädchen (EAS-M) im Alter von 9;0 bis 13;11 Jahre vorgesehen und dient der Erfassung und Prognose aggressiven Verhaltens in

verschiedenen Situationen des Alltags. Gleichermaßen soll er Ansatzpunkte zu Maßnahmen pädagogischer und therapeutischer Intervention aufzeigen.

Zur Erfassung des Verhaltens werden Situations-Verhaltens-Ausschnitte durch Bild-Text-Kombinationen im Test simuliert. Der Proband wird mit 22 Items dieses Musters konfrontiert. Es stellt sich wie folgt dar: Der links oben im Bild stehende Textvorspann umfaßt die Vorgeschichte, in die der dargestellte Situationsausschnitt eingebettet ist. In ihr wird deutlich eine vorausgegangene Aggression geschildert bzw. die Situation als frustrierend oder sonstwie beeinträchtigend dargestellt. Der jeweils rechts unten im Bild stehende verbale Reaktionsteil hält demgegenüber stets drei Möglichkeiten offen, die aufgenommene Handlung fortzusetzen. Neben einer sozial erwünschten Reaktion wird eine leicht und eine schwer aggressive Alternative dargestellt. Der Proband hat sich für diejenige zu entscheiden, der sein Verhalten in der dargestellten Situation entsprechen würde. In den Bildern (sehr differenzierte und viele Details beinhaltende Stichzeichnungen in Schwarzweiß, in denen - gegenüber dem PFT etwa - Ausdruckskomponenten sehr deutlich zum Vorschein kommen) sind durchgehend Interaktionen zwischen gleichaltrigen Kindern und Jugendlichen dargestellt. Bei der Entwicklung der Bilder legte man Wert auf eine gewisse Streubreite (zu Hause, in der Freizeit außerhalb zu Hause, in der Schule).

Der EAS wurde im Rahmen der sozialen Lerntheorie der Aggression

(BANDURA) entwickelt. Die theoretische Differenzierung von Aggression sieht eine Unterscheidung von Bezugsobjekten (Fremdperson, Eigenperson, Gegenstand), Verhaltensmodalitäten (verbal, nonverbal, hinterhältig, direkt) und einen Grad der persönlichen Involviertheit (aktiv teilnehmend vs. parteiergreifender Beobachter / passiv duldend) vor, die in die vorgelegten Reaktionsmöglichkeiten systematisch einbezogen wurde.

In Item Nr. 1 des EAS-J z.B. ist eine Gruppe von Jungen abgebildet, die an einer Mauer steht. Ein Junge steht etwas abseits. Im Textvorspann lesen wir: "Michael, Thomas, Helmut, Torsten und ich sind die Bande der "Fäustlinge". Wir treffen uns fast jeden Nachmittag in unserem Versteck. Heute haben wir uns überlegt, daß jeder einen Decknamen braucht. Da Helmut ein Dickwanst ist, hat sich Michael überlegt, ihn "Schweinchen Dick" zu nennen. Helmut wehrt sich gegen diesen Namen." Die Handlungskette kann nun vom Probanden wie folgt fortgesetzt werden: (a) "Ich stimme mit den anderen für den Namen "Schweinchen Dick"." (leichte, direkte Verbalaggression), (b) "Ich sage, daß wir Helmut nicht mehr länger mit diesem Namen ärgern sollten." (direkte, prosoziale Verbalreaktion), (c) "Da wir abgestimmt und die Mehrheit haben, äußere ich, daß Helmut die Gruppe ja verlassen kann, wenn er mit dem Namen nicht einverstanden ist." (schwere, hinterhältige Verbalaggression) Alle Reaktionsmodalitäten sind dabei zusätzlich aus der Perspektive "parteiergreifender Beobachter" heraus entworfen und vom Typ "Fremdperson bezogen".

Bei der Testauswertung wird die auf die oben geschilderte Art und Weise erhobene Information in eine reaktionsbezogene TYLER-Matrix eingetragen, die zu einer therapiezielorientierten Indikationsstellung ebenso wie zur Veränderungsmessung im Rahmen der Evaluation therapeutischer Interventionsprogramme herangezogen werden kann. Daneben erlaubt sie die Ableitung eines Reaktionsprofils, zu dessen Interpretation dem Benutzer jedoch keine Hinweise gegeben werden. Neben der differentiellen ist für die globale Indikationsstellung die Berechnung eines gewichteten Gesamtpunktwertes für aggressives Verhalten vorgesehen.

Der als durchführungs- und auswertungsobjektiv geltende EAS zeichnet sich durch eine hohe Meßgenauigkeit aus. Für beide Formen (-J, -M) liegen die Schätzungen nach CRONBACHs alpha und GUTTMANs lambda um .90. Nach 8 Wochen wurde für den Gesamtpunktwert des EAS-J eine Retestkorrelation von .71 ermittelt. Eine Faktorenanalyse der Items erbrachte für beide Formen jeweils einen Faktor, den die Items in der Größenordnung von .26 bis .73 laden. Zur externen Validierung wurden Korrelationen zu anderen Tests und Außenkriterien (Lehrer-, Elternurteil) berechnet. Dabei konnte zur Aggressionsskala AG des HAPEF-K kein durchgehender Zusammenhang aufgezeigt werden. Im Hinblick auf die Verhaltensdaten (Aggressions-Ratings) ergaben sich jedoch zum Teil signifikante Beziehungen.

Der EAS wurde als objektiver Verhaltenstest konzipiert und nicht zur Erfassung aggressionsrelevanter Dispositionen. Im Rahmen seiner theoretischen Begründung mutet der Versuch, externe Validität

durch Korrelation zur Aggressionsskala des HAPEF-K aufzuzeigen, etwas befremdlich an. Ansonsten kann man sagen, daß die theoretische Ableitung des EAS vergleichsweise überzeugend und die psychometrischen Qualitäten seiner Kennwerte gut sind. Inwieweit das im EAS simulierte Verhalten die direkte Beobachtung unter natürlichen Bedingungen ersetzt, bleibt abzuwarten. Noch nicht ganz gelöst scheint das Problem, daß durch die unmittelbar-direkte Form der Items der Proband zur Abgabe eines sozial erwünschten Reaktionsprofils verleitet wird (vgl. PETERMANN & PETERMANN 1980, S. 72 f.).

WAGNER (1981) legte eine Verhaltensliste vor, die von den Eltern (Müttern) auszufüllen ist, die "HAMBURGER VERHALTENSBEURTEILUNGSLISTE" (HAVEL). Dieser Fragebogen dient der Erfassung kindlicher Verhaltensauffälligkeiten in Abweichung von der statistischen Norm.

Die HAVEL wurde in der gleichen Hamburger Projektgruppe um HOFSTÄTTER und ähnlich wie der HAPEF-K entwickelt, sodaß man eine dahinterstehende konzeptuelle Übereinstimmung in der theoretischen Begründung vermuten kann (die nicht explizit angegeben wird). Die HAVEL dient jedoch im Gegensatz zum HAPEF-K einer systematisierten, symptomatologischen Bestandsaufnahme via Fremdbeurteilung (beim HAPEF-K via Selbstbeurteilung).

In der Testvorform der HAVEL war neben vier weiteren Skalen (für "vegetative Labilität", "Leistungsverhalten", "Introversion" und "mangelndes Beharrungsvermögen") auch eine für "Aggressivität - negative Dominanz" vorgesehen. In der Testendform schließlich

wurde diese ursprünglich 26 Items umfassende Skala auf eine nunmehr nur noch 10 Items umfassende für "Dominanz" zusammengestrichen. Sie wird hier dargestellt, weil verschiedene ihrer Items aggressionsrelevante Aspekte beinhalten (Wutausbrüche, Neid, feindselige Gefühle).

Zu dem Konstrukt "Dominanz" rechnet WAGNER (1981, S. 16) " ... Verhaltensweisen, die durch ihre starke Affektivität bei häufigerem Vorkommen zu erheblichen Problemen im sozialen Verhalten führen." Und: " ... Geltungsdrang, Jähzorn, Egoismus, angeberisches und besserwisserisches Verhalten."

Zu den als Feststellungen formulierten Items (z.B. "Das Kind neigt zu Wutausbrüchen", "Das Kind wird böse auf seine Eltern" etc.) wird eine fünffach abstufbare Zustimmungsreaktion verlangt, die zwischen "immer oder fast immer" und "nie oder fast nie" variieren kann (wie beim HAPEF-K). Durch Addition über alle Items wird ein gewichteter Gesamtpunktwert ermittelt, dessen Höhe zur Merkmalsausprägung jedoch in einem umgekehrten Verhältnis steht.

Die Reliabilität der Skala "Dominanz" kann als gut bezeichnet werden (interne Konsistenz nach dem varianzanalytischen Ansatz von HOYT = .80 in der Haupt- und .82 in der Nachuntersuchung; die Stabilität nach 23 Monaten beträgt .66). Eine Dimensionierungsanalyse der 40 Items der Testendfassung der HAVEL erbrachte eine klare 4-Faktoren-Struktur, die nach VARIMAX-Rotation mit den einzelnen Subskalen ("Dominanz", "vegetative Labilität",

"Gewissenhaftigkeit" und "Arbeitshaltung") übereinstimmte. Zur internen Validität wird weiter angeführt, daß in der Standardisierungsstichprobe Jungen einen auf 0.1 % - Niveau signifikant niedrigeren Mittelwert erzielten als Mädchen. Zur externen Validität wird eine Korrelation von .23 (sig. 1 %) zu einem Außenkriterium für Aggressivität (Lehrerurteil; keine inhaltlichen Angaben zum Erhebungsverfahren) angegeben sowie eine Korrelation von .19 (sig. 1 %) zur Aggressionsskala des HAPEF-K.

Der Wandel in der Benennung der Skala von der Testvor- zur Testendform bringt bereits unmittelbar zum Ausdruck, daß nicht strikt zwischen Dominanz und Aggressivität getrennt wird. Auch dem Versuch, das Konstrukt zu umschreiben, ist kein klareres Verständnis dessen entnehmbar, was mit der Skala erfaßt werden soll, ebenso wenig der Durchsicht ihrer Items. Das Validierungsvorgehen vertieft die Konfusion. Betrachten wir Validierungsdaten im einzelnen. Unter Bezugnahme auf einen "aggressionstheoretischen" Befund (Mädchen werden eher dazu angehalten, ihre Aggression zu zügeln, Jungen dürfen sie eher zeigen) wird die o.g. Mittelwertedifferenz als Beleg für die (Konstrukt-) Validität der Dominanz-Scores interpretiert. Auch bei den weiteren Daten ist die Logik der Beweisführung die, die Validität der Dominanz-Werte an Aggressionskriterien aufzuzeigen. Gehen wir einmal davon aus, das Lehrer-Urteil sei ein valides Kriterium; dann stellt sich immerhin noch die Frage, was mit dem Zusammenhang belegt werden soll. Diese Korrelation wurde nämlich mit der Testvorform der HAVEL ermittelt. Für die Testendform

wurde die Skala um 62 % verkürzt, und das gerade nicht nach Trennschärfegesichtspunkten, die etwa auf einer Korrelation mit einem Außenkriterium beruhten. Die Korrelation zur Aggressionsskala des HAPEF-K ist u.a. schon deshalb nicht interpretierbar, weil völlig unklar ist, was letztere erfaßt. Insgesamt ist somit festzustellen, daß trotz psychometrischer Bemühungen um formale Qualitäten der Meßwerte die HAVEL-Skala "Dominanz" nicht zur Einschätzung kindlicher Aggression herangezogen werden sollte.

PETERMANN & PETERMANN (1978) stellten ein Rating-Verfahren zur Erfassung von Aggression im Kindergartenalter vor, den "BEOBACHTUNGSBOGEN FÜR AGGRESSIVES VERHALTEN" (BAV).

Wie der EAS, so wurde auch der BAV vor dem Hintergrund der sozialen Lerntheorie der Aggression (BANDURA) entwickelt. Er umfaßt insgesamt 12 Items (Kategorien), von denen allerdings 2 für Dimensionen sozial-erwünschten, prosozialen Verhaltens vorgesehen sind. Speziell dient der BAV einer systematisierten Erfassung kindlicher Aggression nach folgenden dimensionalen Gliederungsgesichtspunkten: verbal (1.1) vs. nonverbal (1.2), offen-gezeigt (2.1) vs. hinterhältig-verdeckt (2.2), Zielobjekt eigene Person (3.1) vs. fremde Person(en) (3.2) vs. Sachen (3.3) und passiv-erfahrend (4.1) vs. aktiv-bewirkend (4.2). Davon sind Itemauswahl und -formulierung logisch abgeleitet, und so lautet z.B. die dritte Kategorie des BAV: "Anschreien, anbrüllen und beschimpfen von Personen." Sie hat den Dimensionsvektor (1.1 2.1 3.2 4.2).

Die Einschätzung auf den einzelnen Kategorien erfolgt jeweils über eine 9-stufige Intensitätsskala von 1 (nicht ausgeprägt) bis 9 (sehr stark ausgeprägt). Der BAV ist nicht zur Ableitung eines Gesamtpunktwertes für Aggression vorgesehen. Es werden lediglich die Ausprägungsgrade auf den einzelnen Kategorien festgehalten (sog. "System von reduktiven Einschätzungen").

Nach einer längeren Einarbeitungsphase in die Handhabung des Instrumentes kann nach den Autoren eine Interraterübereinstimmung von etwa 80 % erreicht werden. Das Verfahren ist augenscheinlich (face) und nach Expertenurteil valide.

Sieht man einmal von der ungeprüften Dimensionalität des BAV ab, so ist vor allem seine unmittelbar einleuchtende Gültigkeit für Aggression sowie - gegenüber dem EAS - die unproblematische Auswertung hervorzuheben. Dennoch ist auch die direkte Beobachtung des Verhaltens als methodischer Zugang zur Erfassung von Aggression nicht ganz unproblematisch, und es wären daher genauere Angaben zu wünschen, die die Brauchbarkeit des BAV belegen (zu den methodischen Problemen vgl. z.B. KOECK & STRUBE 1977).

Zur Erfassung der Wahrnehmung von Gewalt benutzte MOORE (1966) ein Stereoskop. MOORE untersuchte damit bei Kindern und Jugendlichen (Jungen und Mädchen im Alter zwischen 8 und 18 Jahren) differentielle Effekte der Geschlechtsrolle und des Alters auf die o.g. Variable.

Bei der von MOORE benutzten Technik wird einem Auge kurzzeitig (0.5 sek.) eine Gewaltdarstellung, dem anderen - parallel dazu - eine neutrale Szene angeboten. Auf diese Weise werden insgesamt 6 Bildpaare appliziert. Danach erfolgt ein zweiter Durchgang mit Seitenvertauschung des jeweils "aggressiven" und neutralen Bildes.

Auf dem einen Bild des ersten Paares etwa ist ein Briefträger zu sehen, auf dem anderen ein Mann mit einem Messer im Rücken. Bei den Bildern handelt es sich um stilisierte Handzeichnungen in Schwarzweiß. Unmittelbar nach Vorgabe eines Bildpaares soll der Proband genau schildern, was er gesehen hat. Ist er dazu nicht in der Lage, kann die Vorgabe wiederholt werden.

Abgestuft nach dem Grad der Deutlichkeit, mit der der Proband das aggressive Bild beschreibt, wird ihm ein Wert von 2 oder 1 zugeordnet. Beschreibt er klar die neutrale Alternative oder kommt es in seiner Beschreibung zu einer Kontamination der beiden Bildinhalte ohne erkennbare Anzeichen von Gewalt, wird ihm ein Wert von 0 beigemessen. Auf diese Art und Weise wird für jeden Probanden über alle Bildpaare und Durchgänge hinweg ein sog. "violence score" aufaddiert.

Als eine Art "Lügentest" werden zum Schluß als 13. Bildpaar zwei identische Bilder angeboten, die Aggression überdeutlich ("blatant") darstellen. Für Probanden, die diesen Test nicht bestehen, wird kein Wert errechnet. Die im Grunde unabdingbare Sehschärfenkontrolle

wurde von MOORE nur sehr oberflächlich durchgeführt. Probanden z.B., die ihre Brille vergessen hatten, wurden aus der Untersuchung ausgesondert.

Zur Übereinstimmung zwischen zwei unabhängigen Auswertern gibt MOORE eine Produkt-Moment-Korrelation von .98 an.

MOORE konnte mit seiner tachistoskopischen Technik zwei statistisch bedeutsame Befunde erbringen. Jungen erreichen durchgehend höhere Werte als Mädchen, und mit zunehmendem Alter steigen die Werte bei beiden Geschlechtern gleichermaßen an. Darin spiegelt sich nach MOORE eine durch Geschlechtsrolle und Alter modifizierte Sensitivität für die Wahrnehmung von Gewalt in der Umwelt wider. MOORE, der diese Befunde im Rahmen der sozialen Lerntheorie der Aggression (BANDURA & WALTERS 1963) interpretiert, hält es für möglich, daß darunter sowohl eine Wahrnehmungs- als auch eine Verhaltensbereitschaft verstanden werden kann. Es sei theoretisch plausibel, daß mit zunehmendem Alter die Bereitschaft zu (sozial gebilligten) Formen instrumenteller Aggression ansteige.

Wie schon beim RCT oder auch beim FHT geht es hier um eine spezielle Prüfung von Wahrnehmungsinhalten und deren diagnostische Nutzbarmachung. Im Gegensatz zu erstgenannten Verfahren werden diesmal jedoch aggressive Inhalte objektiv vorgegeben und nicht vom Probanden aufgrund einer phantasiegestützten Deutungs- und Interpretationsleistung konstruiert. Es wäre jedoch naiv, diese Technik bereits aufgrund dessen als einen objektiven Test zu

betrachten. Dies würde voraussetzen, daß es dem Probanden grundsätzlich nicht möglich ist, die Zielrichtung des Verfahrens zu durchschauen und auf das, was er berichtet, bewußt Einfluß zu nehmen. Vielmehr muß man davon ausgehen, daß auch der Bericht über das, was man gesehen hat, ein motivierter Prozeß ist, in dem man - je nach kognitiver Einordnung der Untersuchungssituation - Persönliches in einer ganz bestimmten Art und Weise zum Ausdruck bringen oder dies gerade verhindern kann. Am "Lügentest" ist diesbezüglich zu kritisieren, daß er nicht zu einem Zeitpunkt davor vorgegeben wird, wo man eine potentielle Sensitivierung des Probanden hinsichtlich der diagnostischen Zielrichtung noch am ehesten ausschließen könnte; zudem ist er eh nicht sonderlich überzeugend. Durch eine eklatante Darstellung von Aggression erreicht man natürlich, daß es jedem potentiellen Simulanten ohnehin klar wird, daß er gerade unter diesen Umständen nicht lügen sollte. Der Anreiz zum Lügen ist doch offensichtlich in einer mehrdeutigen Situation größer.

Wie läßt sich nun die durch die Technik erfaßte Sensitivität für die Wahrnehmung von Gewalt diagnostisch interpretieren? Aus den von MOORE selbst mitgeteilten Angaben geht das nicht klar hervor. Seine theoretische Spekulation dazu ist etwas pauschal und fußt nicht etwa auf gesicherten empirischen Befunden (vielmehr sollten seine Ergebnisse ja zur Befundlage beitragen). MOORE zitiert in diesem Zusammenhang eine Untersuchung von TOCH & SCHULTE (1961), nach deren Ergebnis etwa Probanden mit einer 3-jährigen Ausbildung im Verwaltungsdienst der Polizei mit dieser Technik (gleiche Bildpaare) einen

signifikant höheren Wert erreichen als eine Gruppe von Probanden, die diesen Ausbildungsgang gerade erst begonnen hat. Für die berufliche Karriere eines Polizisten ist es natürlich von erheblichem instrumentellen Nutzen, eine hohe Sensitivität für objektive Reize zu besitzen, die auf Aggression verweisen. Zusammen mit den von MOORE selbst mitgeteilten Daten würde das auf eine erhöhte Wahrnehmungsbereitschaft hindeuten, ohne dabei jedoch Rückschlüsse auf die zugrundeliegende Motivation zuzulassen. Wahrscheinlich ist in diese Richtung nach dem Gültigkeitsbereich der Technik zu suchen. SARBIN & WENK (1969; zit. n. SELG 1974, S. 76 f.) jedenfalls konnten mit einer vergleichbaren Technik nicht zwischen gewalttätigen und nicht-gewalttätigen Jugendlichen unterscheiden.

Die Reihe der untersuchten Verfahren ließe sich noch fortsetzen. Wir lassen es jedoch dabei bewenden, da wesentlich neuartige Erkenntnisse für den hier verfolgen Zweck nicht resultieren würden und fassen das bislang Gesagte noch einmal kurz zusammen.

3.4 Zusammenfassung

Aggressionsdiagnostik wurde zunächst unter der Perspektive betrieben, ein Abbild innerpsychischer Gegebenheiten zu erzeugen, die man als spezifisch verursachende Bedingungen (Aggressivität) mit Aggression direkt in Beziehung setzte. Die theoretische Patenschaft gewährleisteten das psychodynamische Modell einerseits und das Modell der Persönlichkeitszüge (traits) andererseits.

Die meisten der am psychodynamischen Modell orientierten, projektiven Tests zur Erfassung von (insbesondere feindselig-motivierter) Aggressivität bei Kindern und Jugendlichen konnten jedoch ihren Anspruch bis heute nicht überzeugend einlösen und weisen in der Regel entscheidende psychometrische Mängel auf. Der theoretisch-spekulativ aufgezogene Gültigkeitsnachweis dieser Verfahren stützt sich auf besondere Fähigkeiten des klinisch erfahrenen Benutzers, ohne daß er diese genauer zu spezifizieren vermöchte. Eine Ausnahme stellt eine sorgfältige Entwicklung von LESSER, der L-TAT, dar.

Die auf traitorientierten Persönlichkeitskonzepten beruhenden Fragebögen zur Erfassung von Aggressivität sind zwar in der Regel psychometrisch besser begründet als die o.g. Entfaltungsverfahren und nähern sich insofern der Forderung objektiver Messung, als damit ein sehr viel enger gezogener Interpretationsspielraum für den Benutzer einhergeht. Man fragt sich jedoch aufgrund ihres teilweise - zumindest für Aggression - diffusen theoretischen Überbaus, wozu sie überhaupt entwickelt sind. Diese Frage läßt sich weder intuitiv

durch Inspektion der Items beantworten, noch ist ihren empirisch geführten Gültigkeitsnachweisen eine Antwort darauf zu entnehmen.

Mit zunehmendem Einfluß situationszentrierter Persönlichkeitsmodelle ging eine Akzentverschiebung in der diagnostischen Aufgabenstellung einher. Die sozialen Lerntheorien, bevorzugt an sichtbarem Verhalten und den damit einhergehenden situativen Bedingungen interessiert, legten die Auffassung nahe, die Prognose von Aggression direkt aus letzteren abzuleiten, ohne den Umweg über erschlossene Persönlichkeitszüge und -faktoren zu gehen. Das einzige zur Zeit in Deutschland erhältliche, verhaltenstheoretisch fundierte Aggressions-Meßverfahren in Form eines Persönlichkeitsfragebogens, der EAS, stellt eine etwas unhandliche, aber doch geglückte Synthese zwischen (aggressions-) theoretischen und psychometrischen Erfordernissen dar. Sein Gültigkeitsbereich ist allerdings als eingeschränkt zu betrachten, und es ist fraglich, wieweit er sublime Formen von Aggression (und insbesondere feindseliger) zu erfassen gestattet als auch solche, die über den reaktiven Formenkreis hinausgehen. Letzteres sollte man jedoch nicht unbedingt als Nachteil mißverstehen.

Die apparativen Techniken, von denen wir eine vorgestellt haben, könnten ihren Gültigkeitsbereich als abhängige Variablen (insbesondere zu Kontrollzwecken für die Herstellung experimenteller Bedingungen wie z.B. Ärger) in (aggressions-) theoriebezogenen Experimenten finden und sind dort auch schon mit Erfolg eingesetzt worden (z. B. ZUCKLEY 1978). Zu diagnostischen Zwecken sollten sie vorerst noch nicht herangezogen werden.

4. Problemstellung

Ausgangspunkt der vorliegenden Arbeit war das Bedürfnis, über ein theoriebezogenes Instrument zur Messung des Aggressionsmotivs und der Aggressions-Hemmung bei Kindern zu verfügen. Neben der Theorieableitung sollte das Instrument psychometrischen Standards genügen, wie sie im Rahmen der sog. "klassischen" Testtheorie formuliert worden sind (vgl. z.B. LIENERT 1969; DIETERICH 1973; REY 1977).

Zur Formulierung dieses Bedürfnisses war es nicht ganz zufällig gekommen. Vielmehr standen zwei bis dahin getrennt verlaufene Forschungsrichtungen im Hintergrund, die wir nun fruchtbar miteinander verknüpfen wollten.

Zum einen hatte KORNADT (1982 a, b) im Zusammenhang mit der Entwicklung einer Motivationstheorie der Aggression einen engdimensionierten Aggressions-TAT (A-TAT) für männliche Jugendliche und Erwachsene erstellt, der sich in der empirischen Erprobung bislang als ein (konstrukt-) valides Instrument zur Erfassung individueller Differenzen im Aggressionsmotiv und in der Aggressions-Hemmung erwies (KORNADT 1982 a; ZUMKLEY 1984 a, b).

Zum anderen war - in etwa dem gleichen Zeitraum - im Bereich der Leistungsmotiv-Messung ein sog. "semiprojektiver" (HECKHAUSEN 1980, S. 259) Ansatz der Testkonstruktion (GITTER-TECHNIK) entwickelt, erprobt (SCHMALT 1973, 1974, 1976 a, b) und anschließend auch auf den Bereich des Machtmotivs (SCHMALT 1979) übertragen worden, der

sich den bis dahin innerhalb der Leistungsmotiv-Diagnostik vorherrschenden testmethodischen Ansätzen (teils projektive Verfahren, teils Fragebögen) gegenüber in mehrfacher Hinsicht als überlegen erwies.

Unser Anliegen bestand in concreto darin, der Frage nachzugehen, ob es unter Nutzung der Ergebnisse dieser beiden Entwicklungen möglich sein würde, ein semiprojektiv arbeitendes Verfahren nun auch zur Messung feindselig-motivierter Aggressivität von Kindern zu erstellen, das, wie der A-TAT, theoriebezogene, valide Aussagen über individuelle Differenzen in diesem Persönlichkeitsmerkmal erlaubte, ein "Aggressionsmotiv-Gitter" (AM-GITTER).

Neben der Erwartung, damit in einer weiteren, ergänzenden Population über ein theoriekonformes Meßverfahren zu verfügen, verbanden wir mit der Wahl des o.g. Konstruktionsansatzes vor allem die Hoffnung auf methodisch bedingte Vorteile, wie z.B. auf einen Zugewinn an Durchführungs- und Auswertungsobjektivität, der sich in einer prinzipiell möglichen maschinellen (computergestützten) Vorgabe und Auswertung eines nach der genannten Technik entwickelten Verfahrens widerspiegelt, oder auch auf die umfassender als im Rahmen der Entwicklung rein projektiver Verfahren sich anbietenden Möglichkeiten, das Ergebnis der Entwicklung mit psychometrischen Richtlinien und Gütekriterien in Einklang zu bringen.

Aus einer Reihe von Gründen konnte es sich bei diesem Vorhaben jedoch nicht um eine einfache Übertragung der im Leistungsmotiv-Gitter

("LM-GITTER"; SCHMALT 1976 b), eines erstmals nach dieser Technik entwickelten Motiv-Tests, angelegten Gestaltungsprinzipien und damit verknüpften motivationstheoretischen Annahmen auf einen lediglich anderen thematischen Bereich (Aggression) handeln. Das war zunächst schon daraus ersichtlich, daß wir ein Meßverfahren für einen ganz anderen Motivtyp entwickeln wollten, als ihn das Leistungsmotiv darstellt. Das Aggressionsmotiv ist nämlich nicht nur ein sozial tabuiertes (wie z.B. Sexualität), sondern darüber hinaus auch noch antisoziales Motiv, dessen entwicklungspsychologischen Eigengesetzlichkeiten und funktionellen Besonderheiten wir in erster Linie Rechnung zu tragen hatten (vgl. KORNADT, ECKENSBERGER & EMMINGHAUS 1980).

Was dies im einzelnen bedeutet, soll in den folgenden Abschnitten deutlich werden, in denen wir etwas ausführlicher auf die theoretischen Grundlagen eingehen, die den konzeptuellen Rahmen des AM-GITTERs spezifizieren.

5. Grundlagen der Entwicklung des AM-GITTERs

Theoretisches Bezugssystem des AM-GITTERs bildet die Motivationstheorie der Aggression, wie sie durch KORNADT (1974, 1982 a, c, 1984) ausformuliert wurde. Wir wollen sie in diesem Zusammenhang allerdings nur in der Ausführlichkeit darstellen, wie es zum Verständnis des hier vorgestellten Verfahrens unbedingt geboten scheint.

Methode der Wahl zur Entwicklung des AM-GITTERs ist die GITTER-TECHNIK (GT), wie sie in den Arbeiten von SCHMALT (1973, 1974, 1976 a, b, 1979) dargestellt wurde. Auch sie wollen wir nur in ihren Grundzügen aufzeigen und dabei auf einige ihrer kritischen Punkte aufmerksam machen.

5.1 Motivationstheorie der Aggression

Ähnlich wie DODGE (1982) betrachtet KORNADT Aggression als einen Handlungs-Entscheidungs-Prozeß, dem jedoch - anders als bei DODGE - ein spezifisches Motiv (-system) zugrunde liegt. Dieser Denkansatz ergibt sich aus der Überlegung, aggressives Handeln im Rahmen einer allgemeinen Theorie der Motivation und damit aus der gleichen Perspektive zu untersuchen, wie das zuvor schon (mit Erfolg) z.B. für Leistungs- oder Machthandeln getan wurde (vgl. HECKHAUSEN 1980).

In der von KORNADT nun auf den thematischen Bereich "Aggression"

zur Anwendung gebrachten "Motivationspsychologischen Linie" der persönlichkeitstheoretischen Motivationsforschung (HECKHAUSEN 1980, S. 58 ff.), die als eine Theorie der Wechselwirkung zwischen Person- und Situationsfaktoren gekennzeichnet werden kann, versteht man (z.B. McCLELLAND, ATKINSON, HECKHAUSEN, FUCHS; vgl. KORNADT 1984) allgemein unter einem Motiv ein verhaltenssteuerndes System auf affektiver Grundlage.

Worin bestehen nun die affektive Grundlage des Aggressionsmotivs einerseits und dessen handlungswirksame Funktion andererseits? Treten Behinderung, Bedrohung, Schmerz oder Verletzung eigener Belange ein, so stellt dies eine Frustration dar und ist mit der Aktivierung von Ärger (teils angeboren, teils erlernt) gekoppelt. Ergibt sich nun als Resultat der kognitiven Beurteilung des Gesamtzusammenhangs der Situation, in der die o.g. Emotion auftritt (bedingt z.B. durch die Zuschreibung von Handlungsabsicht, Verantwortlichkeit für die Handlung, durch Bewertung des bedrohten/verletzten Gutes u.ä.), daß sie in gewissem Sinne eine berechtigte Grundlage für Ärger darstellt, so führt dies zu einer endgültigen Ärgeraktivierung und damit Aktivierung des überdauernden Aggressionsmotivs. Ziel des Aggressionsmotivs ist nun die Herstellung eines spezifischen Affektwandels in der Form der Ärgerreduktion. Dieser Affektwandel stellt sich nun mit dem subjektiven Erkennen ein, daß man den Urheber der Frustration direkt oder indirekt in irgendeiner Form geschädigt, beeinträchtigt oder verletzt hat und das von diesem auch so empfunden wird. Der thematische

Person-Umwelt-Bezug oder Gegenstandsbereich des Motivs ist somit "Verletzung, Beeinträchtigung, gewaltsame Beseitigung von Frustationsquellen" (KORNADT 1982 c, S. 96) mit dem Ziel der Herstellung des o.g. Affektwandels.
Handlungswirksam ist demzufolge der als Folge des Handlungsergebnisses von Aggression antizipierte "positive" Affektwechsel in der Form von Ärger-Reduktion (ähnlich wie im Leistungsmotiv der als Folge der Leistungszielerreichung antizipierte "positive" Affektwechsel zum Leistungsstolz).

Es wird nun weiter angenommen, daß dieser antizipierte Affektwechsel unter objektiv gleichen situativen Bedingungen auf unterschiedliche Personen einen unterschiedlich hohen Anreiz ausübt, oder andersherum formuliert, daß individuelle Unterschiede in den "Wertungsvoreingenommenheiten" (HECKHAUSEN) hinsichtlich dieses Affektwechsels bestehen. Diese unterschiedlichen Wertungsvoreingenommenheiten werden mit dem Begriff "Aggressionsmotiv" umschrieben, der damit in dieser Form die Konzeptualisierung individueller Differenzen im Aggressionsmotiv gestattet.

Neben dem Aggressionsmotiv wird als weiteres, auf die gleiche Inhaltsklasse von Handlungen bezogenes Motiv die "Aggressions-Hemmung" angesehen. Affektive Grundlage der Aggressions-Hemmung sind vor allem Furcht vor Gegenaggression und Strafe für Aggression einerseits und Scham und Schuldgefühle wegen der eigenen Aggression andererseits. Handlungswirksam ist hier der "negative" Affektwechsel in der Form der Antizipation von Schuld oder Scham oder auch der positive in

der Form antizipierter Reduktion von Furcht vor Gegenaggression oder Strafangst. Die Verhältnisse sind hier jedoch sehr kompliziert, da nicht nur gleichzeitig unterschiedlich gerichtete Affektwechsel antizipiert werden können sondern sich deren Richtung zusätzlich in Abhängigkeit vom Auftretenszeitpunkt der Hemmung (vor oder nach Aggression) noch einmal ganz unterschiedlich darstellt.
Wie im Aggressionsmotiv, so werden auch hier individuelle Differenzen angenommen.

Mit dem Begriff "antisoziales Motiv" wurde bereits zum Ausdruck gebracht, daß das Aggressionsmotiv einer sozialen Wertung unterliegt. Demgegenüber muß man die Aggressions-Hemmung als ein "prosoziales Motiv" bezeichnen, da ihr Ziel die Verhinderung der Folgen von Aggression oder deren Wiedergutmachung ist.

Aggressionsmotiv und Aggressions-Hemmung bilden nun das Aggressionsmotiv-System. Nach Aktivierung greift dieses System in die Regulation von Handlungs-Entscheidungs-Prozessen ein, und zwar etwa wie folgt:
Unter dem Einfluß des "redintegrierten" Aggressionsmotivs kommt es neben der Ausbildung aggressiver Ziele zur Planung zieldienlicher Handlungen, zur Aktualisierung zweckdienlicher Handlungsmuster sowie zur Vorwegnahme möglicher weiterer Folgen der Zielerreichung; unter Einfluß der Aggressions-Hemmung kommt es zu vergleichbaren Prozessen, deren Zielrichtung jedoch auf die Verhinderung von Aggression verweist. Daneben werden - gemäß dem

Risikowahl-Modell von ATKINSON - hinsichtlich beider Komponenten berücksichtigt (a) die Wahrscheinlichkeit, unter den gegebenen Handlungsmöglichkeiten, die die Situation zuläßt, Erfolg zu haben sowie (b) der Anreiz des für beide Komponenten definierten Erfolgs.

Resultiert nun aus diesen Entwurfs- und Abwägungsprozessen eine positive Gesamtresultante (die Motivation zur Aggressionshandlung ist eine Funktion der Differenz der insgesamt auf Ausdruck und der insgesamt auf Hemmung zielenden Teilprozesse, die untereinander multiplikativ zu einer auf Ausdruck einerseits und Hemmung andererseits gerichteten Teilmotivierung zusammengefaßt sind), so setzt nach einem "Quasi-Entschluß" (Vorsatzbildung) und einem in Einzelheiten noch nicht näher geklärten Willensakt die Handlung (Aggression) ein.

Nach deren Ausführung findet erneut eine Bestandsaufnahme und kognitive Bewertung der nunmehr entstandenen Lage statt, und bei - gemessen an der Erwartungsnorm eines jeweils herzustellenden Schädigungsgrades - "positivem" Ergebnis kommt es zur Desaktivierung der Aggressionsmotivation (Katharsis) und Einstellung der Aggression.

Zum Schluß soll noch auf zwei Punkte hingewiesen werden, die theoretisch bedeutsam sind. (a) Die Theorie befaßt sich ausschließlich mit einer sog. "feindseligen" Aggression (vgl. Abschnitt 2.1); neben der oben dargestellten "Ärger-Aggression" noch mit einer sog. "lustvollen Aggression", für die jedoch andersgeartete Aktivierungsbedingungen angenommen werden. (b) Aggressives Handeln wird danach

bemessen, inwieweit es dazu entworfen wurde, den o.g. Affektwandel bewußt und absichtlich herbeizuführen. Sein in diesem Sinne instrumenteller Wert hängt keineswegs davon ab, inwieweit es nach außen (phänotypisch) erkennbare Aggressionsmerkmale aufweist. So kann z.B. auch gerade die Unterlassung einer Handlung im Sinne einer aggressiven Motivierung zweckdienlich sein.

Zusammenfassend kann man daher sagen, daß die Motivationstheorie der Aggression, von der wir hier ausgehen, von vornherein auf einen sehr viel enger umgrenzten Ausschnitt des Handelns Bezug nimmt, als dies etwa die lerntheoretischen Ansätze tun. Daneben geht sie von individuellen Differenzen aus, die zunächst in der Bewertung der Situation als überhaupt relevant für Aggression, sodann in der Handlungsplanung, Handlungsausführung und Handlungszielsetzung eine wesentliche Rolle spielen und das gerade im Zusammenwirken mit den jeweils vorliegenden Bedingungen der Situation. Der Erfassung individueller Differenzen im Aggressionsmotiv und in der Aggressions-Hemmung kommt somit eine entscheidende Bedeutung zu. Mit ihrer Hilfe soll verstehbar werden, wie die Information, die eine Situation bereithält, in individueller Weise aufgenommen und weiterverarbeitet wird, um auf der Ebene des Aggressionshandelns zur Auswirkung zu kommen.

5.2 Gitter-Technik

Die Gitter-Technik (GT), verstanden als ein methodischer Neuzugang zur Entwicklung und Gestaltung eines Motiv-Tests, geht auf eine Reihe von Einzel- und Gemeinschaftspublikationen auf dem Gebiet der Diagnostik des Leistungsmotivs zurück. Die Veröffentlichungen dazu erschienen mit dem Beginn der 70er Jahre (FISCH et al. 1970; FISCH 1973, 1974; SCHMALT 1973, 1974, 1976 a, b, 1979), die Diskussion darüber befindet sich noch im Gang (z.B. HALISCH 1982, RÖSLER et al. 1982).

Bei einem nach der GT gestalteten Motiv-Test handelt es sich um ein sog. "semiprojektives" Verfahren (vgl. HECKHAUSEN 1980, S. 261). Das erinnert unmittelbar an den als ebenso charakterisierten PFT (vgl. Abschnitt 3.3), dem als testmethodisches Konzept der Gedanke zugrundeliegt, speziell nach einer Theorie gestaltete Bild-Wort-Kombinationen zur Erfassung hypothetischer Konstrukta zu benutzen, die in ihrem Rahmen definiert und begründet sind.

Der semiprojektive Charakter eines Motiv-Gitters kommt nun dadurch zustande, daß zu einer Reihe thematisch strukturierter Bilder, die denen des TAT ähneln (vgl. die Beschreibung des L-TAT in Abschnitt 3.3), nicht mehr Geschichten zu erfinden sind, die hinterher auf ihre motivrelevanten Charakteristika hin ausgewertet werden, sondern die für die einzelnen Komponenten des Motivs diagnostisch bedeutsamen Aspekte in die Form kurzer Aussagen

gegossen und mit den Bildern "gitterwerkartig" zu einer Situations-Aussagen-Matrix verknüpft werden.

Der Proband hat demzufolge (nur noch) die Aufgabe, die einzelnen Situations-Aussagen-Kombinationen der Reihe nach durchzugehen und jeweils die Überlegung anzustellen, ob man die Aussage gewissermaßen als eine Art "Überschrift" einer Geschichte betrachten kann, die einem zu dem betreffenden Bild einfallen könnte. Das Ergebnis seines Nachdenkens gibt er anschließend als JA/NEIN-Urteil durch einfaches Ankreuzen zur Kenntnis. Auf diese Weise gestaltet er die Matrix zu einem hochindividuellen ("idiosynkratischen") Pattern aus, das Rückschlüsse darüber erlaubt, in welcher Art und Weise er die einzelnen Situationen motivthematisch erfaßt und ausgestaltet hat.

Die Operationalisierungslogik der GT ist in einer Wechselwirkungstheorie motivierten Handelns sinnvoll, da sie wie diese auf Umweltseite Inhaltsklassen situativer Anreize und auf Personseite darauf bezogene, überdauernde motivationale Dispositionen zu differenzieren sucht. Einerseits entspricht sie damit persönlichkeitstheoretischen Erfordernissen an ein Motiv-Meßverfahren, denen zufolge bereits in seiner konzeptuellen Gestaltung erkenntlich werden sollte, daß Handeln sowohl eine Funktion der Person als auch eine der Umwelt darstellt, andererseits empirischen Befunden, die darauf verweisen, daß die Wechselwirkung zwischen Person und Umwelt eine bedeutsame Quelle interindividueller Varianz im Handeln darstellt.

Durch die spezielle Verknüpfung TAT-ähnlicher Bilder mit verbalen Kommentaren, die man dann lediglich noch auf Kompatibilität mit diesen zu prüfen hatte, hoffte man, einerseits den Gestaltungsspielraum zu erhalten, den der TAT der Entfaltung motivierter Phantasieproduktionen bot, worauf sich ja seine Validität gründete, andererseits aber auch dadurch, daß das Testverhalten nun fragebogenähnlich erfaßt wurde, die damit einhergehenden Vorteile psychometrisch zu nutzen.

Die Gitter-Technik war insbesondere unter der Perspektive entwikkelt worden, mit dem nach ihr gestalteten Leistungsmotiv-Test, dem LM-GITTER (SCHMALT 1976 b), eine situationsbezogene Motivmessung zu ermöglichen. Unter theoretischen Gesichtspunkten waren die Bildsituationen des LM-GITTERs zu einzelnen Situationsbereichen gruppiert worden (z.B. SCHULE, SPORT, MUSIK etc.) und sollten so zur Ermittlung situationsspezifischer Leistungsmotiv-Kennwerte herangezogen werden können. Insbesondere von FISCH (1973, 1974) war die Auffassung vertreten worden, von der Vorstellung einer generellen, situationsübergreifenden Leistungsmotiviertheit Abstand zu nehmen und sich statt dessen der Erfassung situationsspezifisch getönter Erwartungsemotionen vom Typ der Erfolgszuversicht einerseits und vom Typ der Mißerfolgsbefürchtung andererseits zuzuwenden. Dieser motivdiagnostische Anspruch konnte jedoch bislang durch das LM-GITTER nicht eingelöst werden.

Bei dem Versuch, die theoretisch konzeptualisierte Bereichsdifferenzierung auf Meßwerteebene nachzuweisen, erbrachten faktorielle

Analysen nicht, wie erwartet, Dimensionen, die Inhaltsaspekte von Situationsbündeln beschreiben, sondern zwei Zeitreihen-Verlaufsfaktoren im Sinne einer WOODROW-Struktur (vgl. SCHMALT 1974, S. 262 ff., 1976 a, S. 120 f.; RÖSLER et al. 1982). Dagegen ließ sich in den einzelnen Situationsbereichen die dreifaktorielle Struktur des Gesamt-Gitters stets mehr oder weniger gut replizieren. Der auch von RÖSLER et al. in diesem Zusammenhang zugestimmten Schlußfolgerung, die einzelnen Situationsbereiche lediglich als äquivalente Replikationen im Sinne paralleler Tests aufzufassen und das Gitter insgesamt zu verkürzen, wie sie vorschlagen, muß jedoch entgegengehalten werden, daß zur Erklärung des Zustandekommens der o.g. Struktur Ermüdungsfaktoren herangezogen werden. Somit beschränkt sich die Replikationshypothese notwendigerweise auf den Fall, daß die gesamte Testinformation mit einer einzigen Vorgabe aufgenommen und somit überhaupt die Bedingung hergestellt wird, daß Ermüdung auftritt. Es ist unseres Wissens noch ungeprüft, ob man bei einer zeitlich getrennten Einzelvorgabe der verschiedenen Situationsbereiche zu davon abweichenden Ergebnissen kommt.

U.a. aus o.g. Gründen rät SCHMALT (1976 b, S. 24) vorerst von einer Verwendung situationsspezifischer Kennwerte ab.

6. Bestandteile des AM-GITTERs

Aus dem bislang Gesagten ist deutlich geworden, daß ein nach der GT entwickelter Motiv-Test dem Versuch nahekommt, ein nach TAT-Gesichtspunkten gestaltetes Motivmeßverfahren zu objektivieren. Wie im TAT werden auf der einen Seite Bilder vorgelegt, die die situationsseitigen Bestandteile im Handlungsraum einer Person vergegenständlichen sollen. Dazu werden im TAT die personseitigen Bestandteile von der Person aktiv und im Rahmen eines ausgedehnten Gestaltungsspielraums in der Form einer phantasiegestützten Geschichte konstruiert; in einem Motiv-Gitter dagegen werden sie in Form kleiner Sätze oder kommentierender Aussagen (kurz: "Statements") objektiv vorgegeben. Das schränkt einerseits den Handlungsspielraum der Person in einer gewissen Weise ein. Andererseits reduziert es die subjektiven Anteile bei der Entscheidung, ob ein bestimmter Aspekt (bei der Interpretation einer TAT-Geschichte) als Indikator für eine vorliegende Motivierung (bzw. für ein überdauerndes Motiv) zu werten ist, praktisch auf Null.

Es lassen sich also bei einem nach der GT entwickelten Motiv-Test situationsseitige Bestandteile (Bildsituationen) auf der einen Seite von personseitigen Bestandteilen (Statements) auf der anderen Seite unterscheiden.

Unsere ersten Überlegungen bei der Entwicklung des AM-GITTERs bezogen sich nun (a) auf die Wahl geeigneter Bildsituationen bzw. eines durch sie repräsentierten Handlungsbereichs und (b) auf Möglichkeiten zur Gestaltung motivdiagnostisch relevanter Statements.

6.1 Situationsseitige Bestandteile

Aufgrund der für die endgültige Aktivierung von Ärger konstitutiven Prozesse (z.B. Zuschreibung von feindseliger Verletzungsabsicht, Verantwortlichkeit etc.) ergibt sich, daß sich die situativen Bedingungen im wesentlichen aus den Handlungen des Gegenübers, des Handlungspartners ableiten lassen (vgl. auch HECKHAUSEN 1980, S. 279). Motivationstheoretische Überlegungen legen nun die Auffassung nahe, daß sich über bestimmte Merkmale des Handlungsträgers (Interaktionspartners) situative Bereichsklassen voneinander abheben lassen. Als Parameter der Situation könnte man sich etwa vorstellen das Alter, das Geschlecht, die soziale Rolle des Handlungspartners (vgl. dazu auch die empirischen Befunde bei PARKE & SLABY 1983). Denkbar wären aber auch Bekanntheitsgrad, physisches Erscheinungsbild u.ä. als objektive Merkmale des Handlungspartners. Eine gängige Unterscheidung in der empirischen Aggressionsforschung war diesbezüglich, wenn es um kindliche Aggression ging, die (a) Aggression gegen Erwachsene (v.a. Autoritätspersonen) und die (b) Aggression gegen Gleichaltrige ("peers"; v.a. im Kindergarten und in der Schule untersucht).

Hinsichtlich der Entwicklung des Aggressionsmotivs und der Aggressions-Hemmung kann man nun davon ausgehen, daß dem Situationsbereich "sozialer Umgang mit Gleichaltrigen" eine entscheidende Bedeutung beizumessen ist (vgl. OLWEUS 1974, 1978; MACCOBY 1980; MUSSEN, CONGER & KAGAN 1981; KORNADT 1982 a; PARKE & SLABY 1983).

Da nun der o.g. Situationsbereich in dem Bildersatz des LESSER-TAT (vgl. Abschnitt 3.3) bereits eine vergleichsweise gelungene und auch schon empirisch geprüfte Darstellung gefunden hatte, entschlossen wir uns, ihn als situationsseitigen Bestandteil unseres AM-GITTERs heranzuziehen und auf seine diesbezügliche Verwendbarkeit zu prüfen (zu den Bildern vgl. Anhang I).

Hinsichtlich der o.g. Situationsparameter sind in dem Bilder-Satz des L-TAT insbesondere konstant gehalten:

- das Geschlecht,
- das Alter,
- das physische Erscheinungsbild

und weniger offensichtlich, aber doch wahrscheinlich

- der soziale Status und die soziale Rolle sowie
- der Bekanntheitsgrad des Handlungspartners;

und zwar insofern, als in den Bildern stets die gleichen Jungen dargestellt sind, jedoch in unterschiedlichen Ausschnitten einer komplexeren Handlungskette, die beide Partner (über die Zeit) miteinander verbindet.

Bild Nr. 10 des L-TAT stellt allerdings eine Autoritätsperson (Lehrerin) mit dar und wird aus diesem Grund von uns noch geringfügig modifiziert werden (vgl. Situation Nr. 11 im Anhang I).

6.2 Personseitige Bestandteile

Die Gitter-Technik bietet die Möglichkeit, theoretisch geforderte und/oder empirisch belegte Diskriminatoren motivierten Handelns (wie etwa die aus der Leistungsmotiv-Forschung bekannten typischen Diskrepanzen zwischen Erfolgs- und Mißerfolgsmotivierten hinsichtlich der Anspruchsniveausetzung) als valide und trennscharfe Indikatoren für Motivierungsprozesse in einen Motiv-Test einzubeziehen, was von vornherein zu dessen (theoretischer) Gültigkeit in erheblichem Umfang beizutragen vermag.

Für die Formulierung der Aussagen des AM-GITTERs boten sich daher als theoretische Grundlage die validierten Inhaltskategorien des A-TAT, die bei KORNADT (1982 b) ausführlich beschrieben sind, an. Es handelt sich bei diesen um ein nach theoretischen Gesichtspunkten entwickeltes System zur Erfassung aggressionsmotiv- und aggressionshemmungsrelevanter Aspekte in den Geschichten eines Aggressions-TAT, auf das wir, zusammen mit Überlegungen, wie dieses System im einzelnen benutzt werden soll, im empirischen Teil dieser Arbeit noch genauer eingehen werden.

Die nach den o.g. Gesichtspunkten entwickelten Aussagen sollten dann zusammen mit den Bildern des L-TAT zu einem AM-GITTER verflochten werden und einem Probanden die Möglichkeit bieten, den zugrundeliegenden Situationsbereich aggressionsmotiv- und aggressionshemmungsspezifisch auszugestalten.

6.3 Item

Aus den situationsseitigen einerseits und den personseitigen Bestandteilen andererseits setzen sich die Items des AM-GITTERs zusammen. Die im AM-GITTER anfallenden Daten lassen sich nun unter Rückgriff auf einen schon vergleichsweise "klassischen" Itembegriff, wie er etwa bei TORGERSON (1958, S. 300) zu finden ist, problemlos handhaben. TORGERSON versteht unter einem Item: " ... it will refer to anything that can serve to divide the respondents into two or more mutually exclusive and exhaustive categories." Im Hinblick darauf benutzen wir die einzelne Matrixzelle des GITTERs (Bild-Aussage-Kombination) als Item (-Äquivalent) und treffen noch eine Vorkehrung, die es gestattet, die durch ein Item evozierte Reaktion als Grundlage einer objektiven Aufteilung der Probanden in zwei wechselseitig sich ausschließende Kategorien heranzuziehen. Wir kombinieren deshalb zu jedem Item die Frage: "Paßt der Satz zu dem Bild?" mit den Antwortmöglichkeiten JA und NEIN. Danach sieht ein Item des AM-GITTERs schematisiert folgendermaßen aus (vgl. Abb. 1):

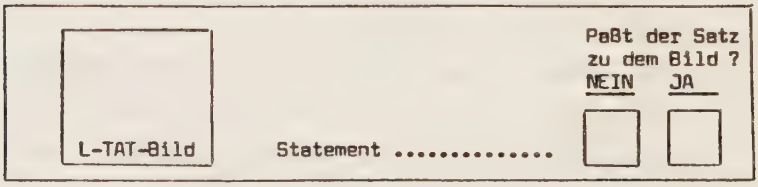


Abb. 1 Schematisierte Darstellung eines Items im AM-Gitter

7. Vorgabe und Funktionsweise des AM-GITTERs

Da wir mit dem AM-GITTER einen TAT-äquivalenten Gültigkeitsbereich anzielen, scheint es uns sinnvoll, die Vorgabe des Verfahrens weitgehend derjenigen anzugleichen, die beim TAT üblich ist. Mit aufgrund dieser Vorgabe sowie der verwendeten Bilder sind wir darüber hinaus der Auffassung, daß sich das Verhalten eines Probanden bei der Bearbeitung des AM-GITTERs weitgehend unter Zuhilfenahme einer Theorie erklären läßt, die Aktivierung und Einfluß motivationaler Prozesse auf den Entwurf und die Gestaltung einer TAT-Geschichte beschreibt.

7.1 Vorgabe

Von kleineren Modifikationen abgesehen haben wir die in den Voruntersuchungen entwickelte Form der Vorgabe des AM-GITTERs beibehalten. Es hat sich insbesondere als nützlich erwiesen, bei der Vorgabe besonderen Wert und besondere Sorgfalt auf die situative Einbettung der Testdurchführung zu legen. Demzufolge unterscheiden wir bei der Vorgabe zwei Phasen: (a) situative Einbettung der Testdurchführung und Aufbau einer geeigneten Testmotivation und (b) Instruktion zur Bearbeitung des AM-GITTERs. Beide Phasen werden nahtlos miteinander verbunden, d.h. dürfen als solche dem Probanden nicht erkenntlich werden.

Die situative Einbettung der Testdurchführung und der daran gekoppelte Aufbau einer zweckdienlichen Testmotivation erfolgt generell

mündlich. Der Erfolg dieser Maßnahme hängt nicht unwesentlich vom Einfühlungsvermögen und der Überzeugungskraft des Testleiters ab. Sie besteht im wesentlichen aus dem Versuch, bei den Kindern eine Art "Expertenmotivation" aufzubauen. Das soll sie vom eigentlichen Zweck der Testdurchführung ablenken, Simulation verhindern und insbesondere auch eine sachbetonte, angstfreie und entspannte Motivationslage anregen. Damit unmittelbar verbunden ist, im weiteren nur solche Kinder mit der Bearbeitung des AM-GITTERs zu beauftragen, bei denen erkenntlich ist, daß das o.g. Ziel auch erreicht wurde. Zur Kontrolle wird daher nach Phase 1 die Frage gestellt, ob jemand unter den Kindern ist, der im weiteren nicht mitmachen möchte; diejenigen, die sich melden, dürfen dann statt dessen etwas anderes tun. Da z.Zt. ohnehin nur Kinder im Gruppenversuch (i.d. Schule) getestet werden dürfen, bei denen im Rahmen der zu erfüllenden Auflagen Freiwilligkeit der Teilnahme weitgehend die Regel ist, kommt es praktisch kaum vor, daß infolge der o.g. Kontrollfrage Kinder ausscheiden.

Zum Aufbau der "Expertenmotivation" wird den Kindern erzählt, daß wir (Institut) eine kleine Bildergeschichte (speziell) für Jungen ihres Alters verfassen möchten. Dazu hätten wir schon einige Bilder ausgewählt und Sätze formuliert, die später unter die Bilder geschrieben werden sollten. Nun seien uns eine ganze Reihe von Sätzen dazu eingefallen, aber wir wüßten noch nicht genau, wie gut sie zu den einzelnen Bildern paßten. Es sei unser Anliegen, dies herauszufinden, und dazu hätten wir uns etwas ganz

Originelles einfallen lassen. Wir hätten uns gedacht, daß sie (Jungen) das eigentlich doch am besten wissen müßten, da sie - unserer Meinung nach - doch als "Experten" (für Jungen dieses Alters) anzusehen seien, und deswegen wollten wir sie einmal schlicht und einfach um ihre Meinung fragen. Nun wird in einem kleinen Dialog abgeklärt, was sie von der "Idee" halten und wer mitmachen bzw. nicht mitmachen möchte.

Nachdem nun festgelegt ist, wer sich als "Experte" für die Befragung zur Verfügung stellt, teilen wir den Kindern die Testformulare aus und sagen ihnen, daß wir uns bereits überlegt hätten, wie sie ihre Arbeit als Experten am besten und ohne Schwierigkeiten durchführen könnten. Dazu hätten wir ihnen eine Arbeitsanleitung auf das Kopfblatt des "Bilderheftes" geschrieben, die wir nun gemeinsam durchlesen und besprechen wollten.

Es folgt nun Phase 2 der Vorgabe, eine TAT-analoge Instruktion zur Bearbeitung des AM-GITTERs. Die Kinder lesen dazu (eines darf laut vorlesen, die anderen lesen still mit) folgenden Text auf der Kopfseite des Testformulars:

WAS DU TUN SOLLST

Du hast ein kleines Heft bekommen. Auf jeder Seite davon ist links ein Bild zu sehen, auf dem zwei Jungen Verschiedenes tun. Rechts daneben stehen kleine Sätze.

Du sollst Dir nun zunächst das Bild einmal genau betrachten. Stell' Dir dabei vor, was die Jungen darauf wohl gerade denken, sich wünschen, was sie tun und wie ihnen dabei zumute ist. Dabei solltest Du Dir ein bißchen Zeit lassen.

Wenn Du damit fertig bist, ließ die Sätze neben dem Bild von oben nach unten aufmerksam durch. Am Ende eines Satzes überlegst Du Dir jeweils, ob er Deiner Meinung nach zu dem Bild dazupaßt oder nicht. Wenn Du "JA" meinst, machst Du ein Kreuz in das rechte Kästchen, wenn Du "NEIN" meinst, in das linke Kästchen neben diesem Satz. Wenn Du Dich nicht entscheiden kannst, was Du ankreuzen sollst, so wähle die Antwort, die Dir lieber ist.

Auf diese Weise kreuzt Du neben jedem Satz von oben nach unten Deine Antwort an. Wenn Du mit einer Seite fertig bist, gehst Du zur nächsten über. Es gibt dabei keine richtigen und keine falschen Antworten. Wichtig ist immer nur Deine ganz persönliche Meinung.

Und nun noch ein paar Tips:

Überlege bei den einzelnen Sätzen nicht zu lange.

Wenn Du einmal nicht sicher bist, was ein Wort bedeutet, so wähle die Bedeutung, die Dir als erste eingefallen ist.

Achte auf folgende Aufmerksamkeitsfehler:

Wenn Du mit einer Seite fertig bist, sollten
☐ ☐ keine zwei leeren Kästchen nebeneinander stehen,
☒ ☒ keine zwei vollen Kästchen nebeneinander stehen.

Und nun viel Spaß!

Es wird nun noch einmal nachgefragt, ob alle die Instruktion verstanden haben und ggfls bestehende Unklarheiten beseitigt. Ferner wird darauf hingewiesen, daß bei irgendwelchen Unklarheiten während der Bearbeitung des Bilderheftes sich der Proband direkt an den Testleiter wenden soll, der dann mit Rat und Tat zur Verfügung steht.

7.2 Funktionsweise

Zunächst ist einmal festzuhalten, daß bislang noch keine gesicherte Theorie darüber existiert, wie man sich die Funktionsweise eines nach der GT entwickelten Motiv-Tests vorzustellen hätte. Dennoch scheint es sinnvoll, zumindest in groben Zügen darzustellen, wie man sich diese vorstellen könnte. Damit soll in etwa die Richtung angedeutet werden, in die man das AM-GITTER als Testverfahren methodisch einzuordnen hätte. Der Begriff "semiprojektiv" deutet diese Problematik bereits an. Danach handelt es sich bei einem Motiv-GITTER um ein Verfahren, das zwischen projektiven Verfahren einerseits und Fragebogen andererseits anzusiedeln ist. Die dadurch bedingte "Zwitterstellung" läßt die Frage offen, wieweit es testmethodisch der einen bzw. der anderen Seite zuzurechnen ist. Diese Frage läßt sich wahrscheinlich auch nur im Rahmen umfangreicher Validierungsuntersuchungen schrittweise klären, wenn man einmal davon ausgeht, daß mit den o.g. methodischen Ansätzen überhaupt unterschiedliche Gültigkeitsbereiche korrespondieren (vgl. McCLELLAND 1951, 1980, 1981).

In diesem Abschnitt geht es also konkret darum, Vorstellungen darüber zu entwickeln, was uns berechtigt, einer Testreaktion einen Hinweis auf eine überdauernde Motivkomponente (Aggressionsmotiv, Aggressions-Hemmung) zu entnehmen. Erst danach wird deutlich werden, was es heißt, auf der Grundlage einer bestimmten Auswertung des AM-GITTERs induktiv Schlußfolge-

rungen über innerpsychische Gegebenheiten auf seiten des Probanden zu ziehen.

Durch die Instruktion wird der Proband zunächst einmal aufgefordert, sich ausführlich mit einem L-TAT-Bild zu beschäftigen. Ein TAT-Bild ist nun ein Ausschnitt aus einer sozialen Situation, der sich hinsichtlich ganz unterschiedlicher Motive thematisch auffassen läßt, d.h. individuell verschieden Möglichkeiten zur Befriedigung ganz unterschiedlicher Motive nahelegt. Wir wissen bereits aus der Beschreibung des L-TAT, daß nun ein L-TAT-Bild mit einer bestimmten Wahrscheinlichkeit Möglichkeiten zur Befriedigung aggressiver Handlungsziele nahelegt. Es soll an dieser Stelle nicht näher darauf eingegangen werden, wie man sich dies im Detail vorstellen kann (vgl. dazu die ATKINSON-FUCHS-Theorie zur Funktionsweise thematisch apperzeptiver Verfahren sowie Theorie dazu von HECKHAUSEN; KORNADT & ZUMKLEY 1982, S. 265 ff.). Jedenfalls kann man - darauf aufbauend - zu der weiteren Annahme gelangen, daß es infolgedessen zu einer phantasiegestützten Setzung aggressiver Handlungsziele und Vorwegnahme damit verknüpfter Affektwechsel kommen kann, was man als einen motivierten Prozeß betrachten kann. Dieser Prozeß funktioniert nun im weiteren quasi als Vergleichsmaßstab, wenn es um die Entscheidung geht, festzulegen, ob ein Statement zu dem Bild "paßt". Es wird angenommen, daß es genau dann paßt, wenn es sich funktionell reibungslos (d.h. "sinnvoll") in den mit dem Bild assoziierten Tätigkeitszusammenhang einordnen läßt oder motivzweckdienlich ist. Dann erfolgt eine

JA-Reaktion. Wird es als nicht motivzweckdienlich erlebt, erfolgt eine NEIN-Reaktion.

Bei den oben angestellten Überlegungen handelt es sich um idealisierte Vereinfachungen. Es wird vorausgesetzt, daß zunächst die gesamte Situation, in der der Beurteilungsprozeß stattfindet, motivthematisch neutral ist. Weiter wird vorausgesetzt, daß die JA/NEIN-Reaktion als Folge einer eigenen Motivzielsetzung betrachtet werden kann. Dies braucht jedoch dann nicht der Fall zu sein, wenn der Anregungsgehalt der Bildsituation so hoch ist, daß sich Kompatibilität von Bild und Statement als Folge nicht-eigener, fremder Motivzielsetzungen ergibt. Drittens wird vorausgesetzt, daß das Statement selbst grundsätzlich individuell geeignet ist, überhaupt eine motivzieldienliche Funktion bzw. einen motivrelevanten Aspekt auszudrücken. Weitere Erfordernisse mögen nicht ausgeschlossen sein. Es ist im weiteren jedoch eine Frage konstruktiver Entscheidungen und Maßnahmen, trotz der oben angedeuteten Schwierigkeiten ein Instrument zu erstellen, das Motivmessung nicht nur vorgibt, sondern es auch nachweislich tut. Es sollte sich also, und das ergibt sich unmittelbar aus dem oben Gesagten, in einer JA-Reaktion nicht eine allgemeine Wertschätzung (motiv-) thematischer Inhalte widerspiegeln, wie sie im Sinne generalisierter Voreingenommenheiten etwa mit einem attitüdentheoretischen Konzept vereinbart werden könnte, aber auch nicht eine lediglich erhöhte "Wahrnehmungsbereitschaft", die vornehmlich den Einfluß situativer Faktoren deutlich macht und auf einen lerntheoretischen Bezugsrahmen verweisen würde (vgl. MOORE 1966).

Insbesondere die letztgenannten Punkte verweisen auf die Unabdingbarkeit einer ausgiebigen, eine Motivationsreaktivierung (Redintegration) in vollem Umfang ermöglichenden Beschäftigung des Probanden mit der Bildsituation, bevor er überhaupt das erste Statement liest und beurteilt. Wir können hier im weiteren zunächst nur einmal vermuten, daß sich bei dem Probanden die in einer Bildsituation reaktivierte motivationale Hauptregung als qualitativer Bewertungsmaßstab durchsetzen wird, sodaß wir in der Folge mit einer unabhängigen Erfassung mehrerer (konkurrierender) Motivierungen werden rechnen können. Auf das Aggressionsmotiv und die Aggressions-Hemmung bezogen heißt das, daß der Proband bei der Bearbeitung einer Situation letztendes zu einer schwerpunktmäßig aggressionsmotiv- oder aggressionshemmungsspezifischen Gesamtauffassung gelangt und damit intern die Voraussetzung schafft, beide Komponenten unabhängig (unkorreliert) erfassen zu können.

Zusammengefaßt funktioniert also das AM-GITTER wie folgt: Bei der Apperzeption der Bildsituationen handelt es sich jeweils um einen motivierten Prozeß, an dem beide Komponenten des Aggressionsmotiv-Systems beteiligt sind. Eine der beiden setzt sich schließlich durch und führt zu einer vorwiegend aggressions- oder hemmungszentrierten Gesamtauffassung, die als qualitativer Bewertungsmaßstab in die Beurteilung der Statements eingreift und damit ein spezifisches Muster in der Situations-Aussagen-Matrix (Gitter) bedingt. Aus dem Umfang, in dem die für die beiden Maßstäbe vorgesehenen Indikatoren (Items) bejaht werden, wird induktiv auf den Grad geschlossen, in dem sie als kompatibel mit dem jeweils dargestellten

Situationsausschnitt erlebt werden, was theoretisch mit dem Stärkegrad einer motivspezifischen Wertungsvoreingenommenheit (Stärke der überdauernden Motivkomponente) gleichgesetzt wird. Über alle Bildsituationen hinweg läßt sich auf diese Weise ermitteln, in welchem Umfang der dargestellte Situationsbereich zu einer Aktivierung des Aggressionsmotivs und der Aggressions-Hemmung beiträgt. Die dabei auftretenden individuellen Differenzen werden als Merkmale der Person (Stärke des überdauernden Aggressionsmotivs bzw. der überdauernden Aggressions-Hemmung) verstanden.

II. EMPIRISCHER TEIL

1. Voruntersuchungen zur Entwicklung des AM-GITTERs

1.1 Vorüberlegungen zur Operationalisierung

1.1.1 Bildsituationen

Wir haben bereits dargelegt, daß wir zur Operationalisierung der situationsseitigen Bestandteile des AM-GITTERs die Bildsituationen des L-TAT verwenden wollen. Im Hinblick auf drei Parameter, nämlich Struktur, Ambiguität und thematischer Anregungsgehalt (vgl. SCHMALT 1976 a, S. 87), läßt sich dieser Bildersatz wie folgt charakterisieren:

(a) STRUKTUR: Handlungsträger sind jeweils zwei Jungen vergleichbaren Alters und vergleichbarer physischer Statur. Beider Handlungsanteile sind jeweils aufeinander beziehbar. Es handelt sich jedesmal um die gleichen Jungen, die in unterschiedlichen Situationsausschnitten dargestellt werden.

(b) AMBIGUITÄT: Hinweisreize (cues), die Motive der Affiliation, Dependenz usw. aktivieren können, sind weitgehend ausgeschlossen. Der Bildersatz ist von einer insgesamt niedrigen Ambiguität.

(c) ANREGUNGSGEHALT: Der Bildersatz verfügt über eine abgestufte Anreizwerteverteilung für Aggression (FA), die (nach den Ergebnissen LESSERs) zwischen 7 % und 94 % variiert. Über den Anregungsgehalt für Aggressions-Hemmung (AA) liegt keine Information vor.

Anmerkung: Bildersatz L-TAT s. ANHANG I

Die Bilder von LESSER scheinen somit für unser Vorhaben grundsätzlich geeignet zu sein. Wir möchten jedoch gerne wissen,

- ob sich die von LESSER mitgeteilten Bildcharakteristika in einer vergleichbaren bundesdeutschen Stichprobe replizieren lassen;
- in welchem Umfang der Bildersatz Aggressions-Hemmungs-Thematik aktiviert;
- ob die auf die Bildvorlagen bezogenen Aggressionsgeschichten vorwiegend feindselige oder instrumentelle Aggression beinhalten, was den Mitteilungen LESSERs ebenfalls nicht zu entnehmen ist (je nachdem sind die auf die (unrevidierten) Bilder bezogenen Statements, mit denen ja eine feindselige Motivierung zum Ausdruck gebracht werden soll, unterschiedlich "plausibel");
- welchen Einfluß die auf dem LESSER-Bild Nr. 3 mit dargestellte Autoritätsperson als spezifischer cue für die Aktivierung des Aggressionsmotivs als auch für die Aktivierung der Aggressions-Hemmung besitzt (1);

und ferner:

- wie sich der Anregungsgehalt der Bilder auf die Itemparameter (Mittelwert, Trennschärfe) beider Konstrukte (Aggressionsmotiv, Aggressions-Hemmung) auswirkt;
- ob sich der Bildersatz als unrevidierter Bestandteil eines GITTERs überhaupt zu einer getrennten Erfassung von Aggressions-Hemmung eignet.

Damit waren erste Fragen formuliert, die es im folgenden zu beantworten gilt.

---

1): Anmerkungen in Klammer s. ANHANG III

1.1.2 Aussagen

Personseitig soll unser Meßinstrument in einem Satz von Aussagen verankert werden, die Konstruktcharakter für die beiden Motivkomponenten Aggressionsmotiv und Aggressions-Hemmung besitzen. Wir haben bereits darauf hingewiesen, daß wir zur Ableitung dieser Aussagen auf den Inhaltsschlüssel des A-TAT von KORNADT zurückgreifen wollen.

Bei einer näheren Betrachtung der einzelnen Kategorien kamen wir zu der Auffassung, daß eine Reihe davon zur Ableitung von Aussagen ungeeignet war. Es handelte sich dabei zunächst einmal um die Kategorien, die mehr auf die formalen Aspekte einer TAT-Geschichte Bezug nehmen. Sodann schieden wir auch die Kategorien für (konstruktspezifische) Symbolik aus. Im Aggressions-Hemmungs-Teil des Schlüssels schienen uns im weiteren diejenigen Kategorien zur Formulierung von Aussagen ungeeignet, die in dem Fall vergeben werden, daß zwischen intrinsischer (aggressionsmotivspezifischer) und extrinsischer (situativ-bedingter) Aggressions-Hemmung nicht eindeutig unterschieden werden kann. Wir wollten nämlich diese beiden unterschiedlichen Aspekte der Aggressions-Hemmung in den Hemmungs-Aussagen getrennt darstellen und prüfen.

Im Aggressionsmotiv-Teil des Schlüssels hatten wir zunächst Überlegungen dazu angestellt, ob auch für die "Diebstahl"-Kategorien Aussagen entwickelt werden sollten. Diese Handlungsmodalität schien rein intuitiv - im Vergleich zu körperlicher, verbaler

sowie auch indirekter Aggression - nicht so recht zu den Bildsituationen zu passen. Wir entschieden uns im weiteren dann so, die Entwicklung von Aussagen der Handlungsmodalität "Diebstahl" davon abhängig zu machen, wieweit letztere den Geschichten zu den einzelnen Bildern selbst entnommen werden konnte. Dies sollte im Zusammenhang mit einer Überprüfung der Bildparameter ermittelt werden.

Eine besondere Schwierigkeit resultierte daraus, daß wir davon ausgehen mußten, daß der Bildersatz des L-TAT vordergründig Aggression nahelegt (für den er ja auch skaliert ist), und gerade nicht Aggressions-Hemmung. Das hätte nun zunächst eine Revision der Bildersatzes oder gar die Entwicklung eines völlig neuen und für Aggressions-Hemmung skalierten bedeuten können. Sieht man von diesen sicherlich sehr arbeitsaufwendigen Möglichkeiten einmal ab, so hatte dies zur Folge, daß die aggressionsthematische Spezifität der Hemmungs-Aussagen irgendwie in die Aussagen selbst hineinverlegt werden mußte. Denn während im (A-) TAT eine Hemmungs-Aussage im Kontext anderer, die ihrerseits Aggression zum Ausdruck bringen, ihre thematische Spezifität gewinnt und somit eine erste Voraussetzung für Indikationsrelevanz erfüllt - letztendes also durch die Art ihrer (psycho-) logischen und funktionellen Einflechtung in den gesamten Handlungszusammenhang der Geschichte, wurde in der GT diese Art der "thematischen Fixierung" noch nicht in Betracht gezogen. Wir dachten an dieser Stelle auch an die Vorgabe ganzer Geschichten, die schwerpunktmäßig entweder Aggression oder

Aggressions-Hemmung darstellen, sahen jedoch aus verschiedenen Gründen dann wieder davon ab. Bei der Vorgabe standardisierter Geschichten geht nicht nur die Eigendynamik des Gestaltungsprozesses verloren, sondern es verringern sich auch - bei gegebenem Zeitaufwand - die Möglichkeiten, einen umfangreichen Itempool zu erstellen und unter den verschiedensten Gesichtspunkten durchzuprüfen.

Wir stellten uns die Lösung dieses Problems so vor, daß die in die Hemmungs-Aussagen eingebauten Aggressionsanteile von den Probanden als wertneutral erlebt werden sollten, d.h. sie durften eben nur die Funktion erfüllen, die ihnen von der Theorie her eingeräumt war (thematische Fixierung). Dagegen sollte ein zweiter Bestandteil einer Hemmungs-Aussage deutlich eine "Wertungsvoreingenommenheit" im Sinne des Konstrukts zum Ausdruck bringen. Es schien uns somit relativ schwierig zu sein, die Aggressions-Hemmungs-Aussagen faktoriell von den Aggressionsmotiv-Aussagen abzuheben.

Auf der anderen Seite stellte sich das Problem einer zweifachen Dimensionierung natürlich auch bei den Aussagen für das Aggressionsmotiv. Dort war es jedoch ungleich leichter zu lösen. Einige der Aggressionsmotiv-Kategorien haben von vornherein mehr eindimensionalen (z.B. "instrumentelle Tätigkeit zur Verletzung", I_v; Bsp.: Sie verhauen sich. Hier tritt der Aspekt einer Wertungsvoreingenommenheit in den Hintergrund, der thematische Bezug dagegen in den Vordergrund.), andere mehr zweidimensionalen Charakter

(z.B. "Bedürnis nach Verletzung", B_V; Bsp.: Er möchte ihn verhauen. Hier treten beide Aspekte etwa gleich stark hervor.). Dieses Problem versuchten wir so zu lösen - und das gilt gleichsam für alle Aussagen des GITTERs - daß jede Aussage gewissermaßen eine Musterlösung für die zur Anwendung gebrachte Kategorie darstellen sollte, womit gleichzeitig die Technik zur Ableitung der konstruktthematischen Aussagen beschrieben ist.

Auch die zuletzt angesprochenen Probleme (z.B. unzureichende Dimensionierung von I_V-Aussagen) ergaben sich daraus, daß die Signierungskategorien insgesamt Geschichten-bezogen zur Anwendung kommen und letztere damit auf ganz anderen Voraussetzungen aufbauen, als es die Gitter-Statements tun, die ja als relativ isoliert dastehende Einzelaspekte auf die jeweilige Bildsituation zu beziehen sind.

Die Probleme einer theoretisch angemessenen Dimensionalität von Aussagen mit Konstruktcharakter werden wir an verschiedenen Stellen der Arbeit noch einmal aufgreifen.

Analog dem Vorgehen von SCHMALT beabsichtigten wir, die Aussagen in die Form kleiner Sätze zu kleiden. Jede Aussage sollte mit allen Bildsituationen kombiniert werden. Daher sollten alle Aussagen inhaltlich gleichermaßen sinnvoll und plausibel auf die einzelnen Bildsituationen bezogen werden können, um Wechselwirkungen zwischen Situationen und Aussagen entgegenzusteuern. Zur Formulierung entschieden wir uns für die "ER-" bzw. für die plurale "SIE"-Form. Das schien insofern sinnvoll, als die auf den Bildern

dargestellten Personen männlichen Geschlechts sind und das Verfahren auch nur für Jungen entwickelt werden sollte (2). Im übrigen sollte über den Verbleib einer Aussage im GITTER und damit implizit auch über die beiden o.g. Formulierungsaspekte anhand psychometrischer Kriterien entschieden werden.

1.2 Erste Schritte zur Konstruktion

1.2.1 Experimentelle Prüfung der Bildsituationen

1.2.1.1 Fragestellung

Es gilt zunächst festzustellen, ob zu den Bildern überhaupt Geschichten erzählt werden, die gemäß den Auswertungsrichtlinien und Kategorien des A-TAT im Kern aggressionsrelevant sind, d.h. als "durch Aktivierung des Aggressionsmotiv-Systems bedingt" (KORNADT 1982 b, S. 23) angesehen werden können. Auf diese Weise ist ihre grundsätzliche motivthematische Eignung zu ermitteln.

Etwas detaillierter soll sodann der Frage nachgegangen werden, in welchem Umfang die beiden Teilkomponenten des Motiv-Systems, Aggressionsmotiv und Aggressions-Hemmung, daran beteiligt sind als auch, in welchem Umfang instrumentelle Aggression mit den einzelnen Bildsituationen in Zusammenhang gebracht wird.

Was zu den Bildern sonst noch an spezifischen Thematiken assoziiert wird (Ambiguität), ist aus mehreren Gründen wichtig zu wissen. Zum einen erhoffen wir uns davon Hinweise, die im Hinblick auf die Revision oder das Aussondern einzelner Bilder

relevant sein könnten. Zum anderen wollen wir bei der Formulierung der konstruktthematischen Aussagen die Richtungen möglicher, d.h. bildbezogen zu erwartender, Fehldeutungen abschätzen können, wenn sie vom Probanden als Bestandteile eines GITTERs gelesen und auf die einzelnen Bildsituationen bezogen werden müssen.

Nicht zuletzt wollten wir wissen, wieweit Bild Nr. 3 Aggression gegen Autoritätspersonen hervorruft.

1.2.1.2 Methode

Die Bilder werden gemäß den Durchführungsbedingungen des L-TAT einer vergleichbaren Stichprobe von Jungen im Alter zwischen 9 und 14 Jahren vorgelegt. Die dazu verfaßten Geschichten sollen anschließend das Rohmaterial zur Beantwortung der gestellten Fragen bereitstellen.

Unsere Stichprobe gliedert sich in verschiedene Teilstichproben auf, die wir S1, S2 und S3 nennen wollen. Bei S1 handelt es sich um zwei vierte Klassen einer saarländischen Grundschule (Jungen, 9 - 11 Jahre, n = 32), bei S2 um zwei siebte Klassen einer saarländischen Realschule (Jungen, 13 - 14 Jahre, n = 36) und bei S3 um eine Auswahl von Jungen einer saarländischen Heimsonderschule für Verhaltensgestörte (10 - 14 Jahre, n = 17). Wir lassen S3 vom Leiter der Einrichtung nach folgenden Gesichtpunkten zusammenstellen: normale Schulleistung und Intelligenz, besondere Resistenz gegenüber sozial- und heilpädagogischen Interventionen jeglicher

Art, Aufsässigkeit.

S1 repräsentiert die Population, für die das AM-GITTER entwikkelt werden soll. In S2 erwarten wir detaillierter und differenzierter ausgestaltete Phantasiegeschichten, denen wir in höherem Umfang Anregungen zur sprachlichen Formulierung der Gitter-Statements entnehmen können. In Zusammenfassung mit S1 soll die gepoolte Stichprobe S1/S2 eine den LESSER-Befunden vergleichbare Datengrundlage liefern. S3 ist als Extremgruppe gedacht, bei der uns vor allem interessiert, wieweit sie in ihren Ergebnissen von denen der beiden übrigen abweicht.

Aus verschiedenen Gründen (Vorbeugung gegenüber einer potentiellen Überlastung der Probanden, Steigerung der Ergiebigkeit der Phantasieproduktionen, Herabminderung der Wahrscheinlichkeit eines möglicherweise eintretenden "Serien-Effekts" (vgl. KORNADT & ZUMKLEY 1982, S. 276)) halbieren wir den Bildersatz und legen der einen Hälfte jeder Teilstichprobe die so erzeugte Version A, der anderen die Version B vor, wobei wir darauf achten, daß jeweils nebeneinandersitzende Jungen unterschiedliche Versionen erhalten. Als Halbierungskriterium wählen wir die von LESSER (1958 b) mitgeteilten Anreizwerte für Aggressionsthematik, denen zufolge auf die Version A die Bilder mit den Nummern 2, 4, 6, 8, 10 und auf die Version B die Bilder mit den Nummern 1, 3, 5, 7 und 9 entfallen (LESSER-Numerierung).
Die Bilder werden zu zwei Bilderbögen in DIN A 4 - Format zusammengefaßt. Dazu wird ein vorstrukturierter Antwortbogen gleichen

Formats vorgelegt, auf dem einige Gliederungshinweise dem Probanden helfen sollen, eine möglichst vollständige Geschichte zu schreiben (z.B. Wer sind die Personen auf dem Bild, wie heißen sie?).

Mittels eines projizierten Diapositivs (Bild Nr. 3 des A-TAT, neutral, kein Aggressions-Anreiz; zwei männliche Erwachsene sitzen nebeneinander an einem Tisch) wird die Aufgabenstellung zuvor an der Tafel eingeübt.

Die Untersuchungen finden jeweils vormittags während der offiziellen Unterrichtszeit statt. Lehrpersonen sowie die Mädchen der jeweiligen Klassen sind nicht anwesend, besondere Störungen nicht zu verzeichnen. Die Atmosphäre ist stets gelockert, freundlich und entspannt.

Wir analysieren die Geschichten nach den Auswertungsrichtlinien und Kategorien des A-TAT. Getrennt vermerken wir instrumentelle Aggression und andere Thematiken. Um eine Vergleichbarkeit mit den Befunden LESSERs herzustellen, übernehmen wir dessen 0/1-Scorung. Zur Ermittlung der Anreizwerte stellen wir fest, in wieviel Prozent der Fälle ein Hinweis auf die jeweilige Thematik in den Geschichten festgestellt werden konnte.

1.2.1.3 Ergebnisse

MOTIVTHEMATISCHE EIGNUNG: Die Inspektion der Geschichten ergibt, daß zu allen Bildern sowohl Aggressions- als auch Aggressionshemmungs-Thematik geäußert wird.
Die in den Handlungsabläufen gestaltete Aggression ist überwiegend feindseliger Art und läßt sich weitgehend als frustrationsbedingte Ärger-Aggression typisieren.
In nur wenigen Fällen, die alle auf Bild Nr. 1 beschränkt bleiben, wird in den Geschichten deutlich instrumentelle Aggression erkennbar. Ungleich häufiger jedoch wird mit dieser Bildsituation feindselige Aggression der o.g. Art in Verbindung gebracht, bisweilen auch mit instrumenteller Aggression durchmischt.
Bild Nr. 3 aktiviert vorzugsweise Gleichaltrigen-Aggression, weniger Aggression gegen Autoritätspersonen (letztere insbesondere in S3); dagegen nicht selten Strafe für Aggression durch Autoritätspersonen.

AMBIGUITÄT (3): Alternativ zu Aggression wird meist etwas über "Freundschaft" geschrieben, was sich bei der speziellen Struktur des Bildersatzes auch leicht einsehbar als gegenläufiges Motiv anbietet. Nur vereinzelt werden auch andere Themenbereiche registriert; so z.B. Gefahr und Angst vor Verletzung (Bild Nr. 9), Leistung (Bild Nr. 3). Verschiedene Arten von Konflikt, Langeweile, Unlust als auch reine Bilddeskriptionen runden den Bereich der assoziierten Themen ab. Der Bildersatz ist insgesamt von einer niedrigen Ambiguität.

ANREGUNGSGEHALT: Die Anreizwerte der einzelnen Bilder für das Aggressionsmotiv stimmen insbesondere bei den Grundschülern (S1) recht gut mit den von LESSER mitgeteilten Anreizwerten für Aggressionsthematik (FA) überein. Wir ermitteln im einzelnen eine kontinuierlich abgestufte Anreizwerteverteilung zwischen 100 % und 6 % in S1, zwischen 100 % und 15 % in S2 und zwischen 100 % und 0 % in S3. Für S1/S2 (gepoolt) variieren die entsprechenden Werte zwischen 100 % und 11 %, für S1/S2/S3 zwischen 100 % und 9 %. Damit korrespondieren (wie auch im folgenden) jeweils leichte Verschiebungen in der Rangordnung der einzelnen Situationen.

Mit Ausnahme von Bild Nr. 9 in S1 , bei dem in gleichem Umfang eine Anregung des Aggressionsmotivs wie auch der Aggressions-Hemmung festgestellt wird, finden sich keine Hinweise darauf, daß einzelne Bilder, sei es durch spezifische cues oder durch die Gesamtsituation als komplexen cue, primär Aggressions-Hemmung angeregt wird. Man kann den Bildern insgesamt jedoch eine "sekundäre" Anregungsqualität für Aggressions-Hemmung zusprechen. Das ist so zu verstehen, daß eine aggressionsmotivthematische Deutung der Bildsituation und daran gekoppelte thematische Schwerpunktsetzung in der Geschichte in etwa zwei Drittel aller Fälle (S1/S2/S3) deutlich erkennbar die Darstellung intrinsischer als auch extrinsischer Hemmung von Aggression nach sich zieht, wohingegen (aggressionsmotivspezifische) Hemmung nie lösgelöst von der o.g. Voraussetzung auftritt.

Die Anreizwerte für Aggressions-Hemmung lassen sich ebenfalls kontinuierlich abstufen; es ergibt sich dann zum Teil eine andere Rangordnung der Situationen, als wenn man sie nach den Anreizwerten für das Aggressionsmotiv rangordnet. Die Ergebnisse deuten jedoch darauf hin, daß diese Unterschiede mit zunehmendem Alter der Probanden als auch mit zunehmender Stichprobengröße in den Hintergrund treten.

Im einzelnen variieren die Anreizwerte für Aggressions-Hemmung zwischen 0 % und 40 % in S1, zwischen 10 % und 70 % in S2, zwischen 0 % und 50 % in S3, zwischen 5 % und 54 % in S1/S2 gepoolt sowie zwischen 4 % und 48 % in S1/S2/S3 gepoolt.

1.2.1.4 Diskussion

Die Bilder scheinen zur weiteren Verwendung alle geeignet zu sein. Im Hinblick auf die Ausgestaltung einer damit verknüpften Phantasiegeschichte lassen sie (bis auf Bild Nr. 7) zwei thematisch unterschiedliche Schwerpunktsetzungen zu: eine aggressions- und - alternativ dazu - eine freundschaftsthematische. Eine simultane Aktivierung beider Motivsysteme innerhalb einer Bildsituation ist in diesem Zusammenhang nicht erkennbar. Gleichwohl kommt es jedoch schon einmal vor, daß ein Proband einen die einzelnen Bildsituationen übergreifenden Gesamtzusammenhang zwischen den Geschichten herstellt - was durch die strukturellen Eigenschaften des Bildersatzes ja begünstigt wird - und dann den Deutungswechsel im Sinne der o.g. Wahlmöglichkeiten als konflikthaft erlebt, wie wir in

kurzen, informellen Gesprächen im Anschluß an die Untersuchungen feststellen konnten. Dies scheint jedoch eher die Ausnahme zu sein.

Die Ergebnisse lassen darauf schließen, daß der Bildersatz (aufgrund besonderer Aspekte seiner graphischen Gestaltung) direkt das Aggressionsmotiv und indirekt (über ein aktiviertes Aggressionsmotiv) die Aggressions-Hemmung aktiviert. Motivationstheoretisch stellt o.g. Auslösebedingung für Aggressions-Hemmung den Regelfall dar. Da wir noch nicht wissen, wie sich diese spezielle Anreizcharakteristik für Aggressions-Hemmung in einem Motiv-Gitter darstellt, halten wir es daher für gerechtfertigt, alle Bilder zunächst einmal unrevidiert zu übernehmen, um sie als situationsseitigen Bestandteil eines GITTERs zu erproben.

Aufgrund der Ergebnisse entschließen wir uns, eine Modifikation von Bild Nr. 3 herzustellen. Wir entfernen einfach die dort dargestellte Autoritätsperson, eine Lehrerin, um das Bild auf diese Weise strukturell den übrigen noch mehr anzugleichen. Wir prüfen diese Modifikation (zusätzlich) als Bildsituation Nr. 11, wenn wir die Bilder erstmals als Bestandteile eines GITTERs in Untersuchung geben.

1.2.2 Entwicklung konstruktthematischer Aussagen

1.2.2.1 Wahl der Kategorien

Neben den bereits erwähnten scheiden wir nun auch die der Verhaltensmodalität "Diebstahl" zugeordneten Aggressionsmotiv-Kategorien aus, da wir in den Geschichten zu den LESSER-Bildern so gut wie keine Hinweise für ihre Anwendung gefunden haben. Zudem glauben wir, daß sie in Bezug auf die noch verbleibenden Kategorien von nachrangiger Bedeutung sind.

Mit den restlichen Aggressionsmotiv-Kategorien lassen sich nun folgende Aspekte eines aggressionsmotivierten Handlungsablaufs darstellen:

- Bedürfnis nach Verletzung (Bv),
- dabei auftretende, aggressionsrelevante Gefühle (Gv),
- instrumentelle Tätigkeit zur Verletzung (Iv),
- Erwartung des Aggressionserfolges (Ev),
- Erreichen des aggressiven Handlungsziels (Zv).

Wir unterscheiden innerhalb jeden Aspektes zusätzlich noch einmal, ob dieser mit körpermotorischen (k), verbalen (v) oder höhergeneralisierten Handlungsschemata (a) verknüpft zu sehen ist.

Die nachfolgend dargestellte Matrix stellt noch einmal die Gesichtspunkte zusammen, unter denen wir Aussagen mit Konstruktcharakter für das Aggressionsmotiv (A-STATEMENTS) entwickeln wollen (vgl. Abb. 2).

A-STATEMENTS	_Bv	_Gv	_Iv	_Ev	_Zv
k	kBv	kGv	kIv	kEv	kZv
v	vBv	vGv	vIv	vEv	vZv
a	aBv	aGv	aIv	aEv	aZv

Abb. 2 Kategorienmatrix zur Ableitung der A-STATEMENTS

Desgleichen versuchen wir, folgende Hinweise auf Aggressions-Hemmung (AH) zu operationalisieren:

- Bedürfnis nach Vermeidung von Aggression (B),
- instrumentelle Tätigkeit zur Vermeidung von Aggression (I),
- hemmungsrelevante Gefühle (G)
- Erwartung unangenehmer Folgen von Aggression (E)
- Vereitelung von Aggression (V)
- Strafe für Aggression (S).

Hier unterscheiden wir noch einmal, wie jeder dieser Aspekte in einer ganz typischen Weise für den extrinsischen (s) einerseits und für den intrinsischen (r) Teil der Aggressions-Hemmung andererseits gestaltet werden kann.

Abbildung 3 faßt den Teil des Schlüssels, den wir zur Ableitung der Aussagen mit Konstruktcharakter für die Aggressions-Hemmung (AH-STATEMENTS) heranziehen wollen, noch einmal zusammen.

AH-STATEMENTS	B_	I_	G_	E_	V_	S_
s	Bs	Is	Gs	Es	Vs	Ss
r	Br	Ir	Gr	Er	Vr	Sr

Abb. 3 Kategorienmatrix zur Ableitung der AH-STATEMENTS

Wir erstellen nun eine umfangreiche Liste mit Vorschlägen für die Anwendung der oben dargestellten 27 Kategorien und legen sie, bevor wir eine erste Auswahl treffen, zwei Experten des A-TAT zur Einschätzung vor. (4)

1.2.2.2 Expertenrating

Zwei Experten des A-TAT werden gebeten, pro Kategorie diejenigen Statements zu markieren, die bestimmte Kriterien erfüllen. Dazu gehören neben der eindeutigen Anwendbarkeit der jeweiligen Schlüsselkategorie und der potentiellen Anwendbarkeit auf alle Bildsituationen eine angemessene sprachliche Gestaltung, Kürze, niedriger Anreizwert u. ä. Die AH-Statements sollen außerdem einen neutralen aggressionsthematischen Bezug aufweisen, um die Motivspezifität sicherzustellen, gleichzeitig aber keine normativen Richtlinien zum Ausdruck bringen, die Anreizcharakter für Zustimmung im Sinne sozialer Erwünschtheit besitzen. Zudem sollen sie so unspezifisch formuliert sein, daß sie gleich gut auf verschiedene Modalitäten aggressiven Handelns (körperlich, verbal,

indirekt, symbolisch etc.) bezogen werden können.

Nach der Einschätzung der Aussagen durch die Experten wählen wir aus der Liste 30 A-Statements und 26 AH-Statements aus. Für die einzelnen der 27 zur Entwicklung von Statements herangezogenen Kategorien liegen nun zwischen einem und vier Vorschläge vor (im Durchschnitt zwei), die wir als Bestandteile eines erstmals erstellten GITTERs prüfen wollen.

1.2.3 Entwicklung und Funktion neutraler Aussagen

Neben den zu diagnostischen Zwecken vorgesehenen enthält ein Motiv-Gitter noch einen weiteren Typ von Aussagen, sog. "neutrale" oder "Puffer-Aussagen". Sie dienen in der Hauptsache der Verunklärung der Untersuchungsabsicht bzw. der diagnostischen Zielrichtung des Verfahrens. Damit soll einem GITTER der Anreiz zum Abgeben von Antworten im Sinne sozialer Erwünschtheit entzogen werden. Zudem sollen sie dem Probanden die Möglichkeit bieten, die Anregung einer nicht in der diagnostischen Zielrichtung des Verfahrens liegenden Motivierung überhaupt zum Ausdruck bringen zu können.

In unserem Fall liegt nun ein besonders günstiger Umstand mit dem Sachverhalt vor, daß sich in den einzelnen Bildsituationen alternativ zu Aggression vordergründig eine Motivierung zeigt, von der man annehmen kann, daß ihre Handlungsziele mit denen von Aggression nicht zu vereinbaren sind. Es handelt sich dabei

um Freundschaft, von der man zudem annehmen darf, daß sie eine feindseliger Aggression entgegengesetzte soziale Wertung erfährt. Es liegt deshalb nahe, die neutralen Aussagen des AMGITTERs aus diesem thematischen Bereich heraus zu entwickeln. Damit sollte auch eine Möglichkeit geschaffen werden, den einzelnen Bildsituationen Hinweise auf die Logik ihrer thematischen Strukturierung durch die Probanden zu entnehmen und daran ggfls. Simulation zu erkennen.

Für die erste Testvorform werden aus den restlichen Themenbereichen, die wir den TAT-Geschichten zu den Bildern entnommen haben, noch weitere Puffer-Aussagen entwickelt. In nachfolgenden Testversionen sind dann nur noch Puffer-Aussagen aus dem Bereich Freundschaft enthalten (F-STATEMENTS). Dies u.a. auch deshalb, weil wir an die Ergebnisse bei der Bestimmung ihrer Dimensionalität ganz bestimmte theoretische Erwartungen knüpfen.

Die Liste der entwickelten Aussagen steht im ANHANG II.

1.3 Zusammenstellung und Prüfung einer Testvorform

1.3.1 Vorbefragung

Ziel der Vorbefragung ist es herauszufinden, welche Probleme bei der von uns gewählten Gestaltung des Testmaterials auftreten, ob das Verfahren gemäß der erforderlichen Bedingungen durchgeführt werden kann und mit welchen Testbearbeitungszeiten gerechnet werden muß.

Wir heften nun 4 Bildsituationen zu einem Bilderbogen zusammen. Dazu erstellen wir einen Antwortbogen, der eine kleine Auswahl von Aussagen beinhaltet. Jede Seite des Antwortbogens enthält die gleiche Auswahl von Aussagen in der gleichen Reihenfolge von oben nach unten. Somit wird der gesamte Aussagen-Satz unverändert mit jeder Bildsituation kombiniert.

Wir legen dieses "Mini-GITTER" zwei vierten Grundschulklassen (Jungen, 9 - 11 Jahre) zur Bearbeitung vor. Anschließend führen wir ein explorativ es Interview durch und schneiden es auf Band mit.

Es zeigte sich, daß es bezüglich der Instruktion zur Bearbeitung des Verfahrens als auch hinsichtlich der Bearbeitung selbst zu keinerlei Mißverständnissen und Komplikationen gekommen war. Man konnte daher beides in der gegebenen Form beibehalten.

Die Probanden empfanden es als angenehm, ihre eigene Meinung äußern zu dürfen und - abgesehen von formalen Ankreuzfehlern - nichts falsch machen zu können. Die Bearbeitung des Verfahrens erlebten sie weder als ermüdend noch als anstrengend oder frustrierend.

Es zeigte sich, daß es auch hier vereinzelt zur Ausbildung situationsübergreifender Handlungszusammenhänge kommen und der von Aggression zu Freundschaft alternierende Deutungswechsel in der thematischen Strukturierung der Situationen als konfliktheft empfunden werden konnte.

Für eine ausführliche Instruktion erwiesen sich 15 Minuten als ausreichend. Die Testbearbeitungszeiten variierten - rückgerechnet - zwischen 6 sek. und 16 sek. pro Item.

1.3.2 Testzusammenstellung

Zur Darstellung des Situationsbereichs der Testvorform wählen wir die 10 LESSER-Bilder sowie unsere Modifikation von Bild Nr. 3 als Situation Nr. 11 (vgl. ANHANG I). Die 30 A-Statements und die 26 AH-Statements (vgl. ANHANG II), die wir zur erstmaligen Prüfung als Bestandteile eines GITTERs entwickelt und ausgewählt haben, puffern wir mit weiteren 16 neutralen. Den gesamten Satz von Aussagen teilen wir auf zwei DIN A 4 - Blätter auf. Der erste Teil soll als Version A und der zweite als Version B eines Antwortbogens zur Anwendung kommen. Die beiden Aussagen-Sätze werden in jeweils unveränderter Reihenfolge mit allen Bildern kombiniert und erlauben je nach Auffassung der Bildsituation ein unterschiedliches Ankreuzmuster. Die Bildsituationen werden in einem Bilderbogen zusammengefaßt. Wir verwenden sie in Originalgröße (DIN A 4).

1.3.3 Durchführung der Untersuchung

In die Untersuchung werden 17 vierte Klassen aus 6 saarländischen Grundschulen einbezogen. Untersucht werden jeweils nur die Jungen einer Klasse. Es nehmen alle Schüler teil, die am Tag der Untersuchung anwesend sind. Die Gesamtstichprobe umfaßt 247 Probanden.

Man kann sie als repräsentativ für Jungen im Alter von 9 bis 11 Jahren aus mittelstädtischen Industriegebieten des Saarlandes betrachten.

Die Untersuchungen finden morgens während der offiziellen Unterrichtszeit in separaten Schulräumen statt. Nach Möglichkeit werden alle Klassen der betreffenden Schule simultan getestet. Damit wollen wir verhindern, daß Probanden bereits mit "Vorinformationen" an den Untersuchungen teilnehmen.

Jeweils die eine Hälfte der Probanden erhält zunächst den Antwortbogen A und fährt anschließend mit der Bearbeitung von Bogen B fort; bei der anderen Hälfte ist es umgekehrt. Nebeneinandersitzende Probanden beginnen mit unterschiedlichen Versionen. Dadurch wollen wir einerseits einer möglicherweise auftretenden Tendenz, in den Antwortbogen des Nachbarn hineinzuschauen, entgegensteuern, andererseits potentiell auftretende Ermüdungseffekte, die sich ohne Splitting des Antwortbogens einseitig auf die zweite Itemhälfte niederschlagen könnten, ausbalancieren.

Es bestehen keine Zeitlimits. Nach individuellem Ermessen können die Probanden in den Testräumen eine Pause einlegen. Die einzelnen Untersuchungen dauern im Mittel etwa 4 Stunden. Alle Antwortbögen werden bei Abgabe auf Unvollständigkeit der Bearbeitung und formale Fehler kontrolliert und anschließend verbessert.

Lehrpersonen sind bei den Untersuchungen nicht anwesend, besondere Vorkommnisse nicht zu verzeichnen.

1.3.4 Prüfung der Bildsituationen und Aussagen

1.3.4.1 Fragestellung

Wir wollen die personseitigen (A- und AH-Statements) und situationsseitigen Bestandteile der Testvorform (Bildsituationen) zunächst unter konventionellen psychometrischen Gesichtspunkten (Mittelwert, Trennschärfe) charakterisieren, um erste Aufschlüsse über ihre Verwendbarkeit zu erhalten. Die Überprüfung der Bildsituationen soll für das Aggressionsmotiv und die Aggressions-Hemmung getrennt vorgenommen werden.

Ein weiteres Ziel der Prüfung ist es, die o.g. psychometrischen Aussagen-Charakteristika im Hinblick auf den aggressionsthematischen einerseits und den aggressionshemmungsthematischen Anregungsgehalt der Bildsituationen andererseits zu betrachten. Daraus lassen sich ggfls. Rückschlüsse für eine gezielte Änderung der Bildsituationen ableiten.

Schließlich sollen die Trennschärfen sowohl der Aussagen als auch der Bildsituationen auf statistisch bedeutsame Unterschiede geprüft werden, um nachfolgende Entscheidungen (Veränderungen in den oder Reduktion von Bild- und/oder Aussagenbestandteilen der Testvorform) besser begründet vornehmen zu können.

Bei günstiger Ergebnislage sollen weitere Untersuchungen folgen. Zielvorgaben dafür sind die Bestimmung der Dimensionalität der

Motivindikatoren (Aussagen) und - auf diesen Ergebnissen aufbauend - die Prüfung von Möglichkeiten zur Skalenbildung sowie die Bestimmung der Dimensionalität eines auf einem bestimmten Satz von Skalen beruhenden GITTERs.

1.3.4.2 Methode

Wir zerlegen die gesamte Situations-Aussagen-Matrix der Testvorform in zwei Teile, und zwar in einen Aggressionsmotiv-Teil (Situationen x A-Statements) und in einen Aggressions-Hemmungs-Teil (Situationen x AH-Statements).

Im Aggressionsmotiv-Teil berechnen wir pro Item die Wahrscheinlichkeit einer Zustimmungsreaktion (p_{ja}) als Mittelwert und die Trennschärfe als biseriale Korrelation (r_{bis}). Trennschärfekriterium ist die Summe der über alle Items ausgezählten JA-Antworten. Nun werden die beiden Itemstatistiken pro Situation (Zeile) und Aussage (Spalte der Teilmatrix) arithmetisch gemittelt. Mit den resultierenden Ergebnissen charakterisieren wir Mittelwerte und Trennschärfen der Situationen (jeweils über alle Aussagen) und Aussagen (jeweils über alle Situationen) im Aggressionsmotiv-Teil der Situations-Aussagen-Matrix. Zur Bestimmung von Trennschärfeunterschieden werden 5 % - Konfidenzintervalle für die Parameterschätzungen berechnet und auf Überlappung geprüft.

Ebenso verfahren wir im Aggressions-Hemmungs-Teil der Situations-Aussagen-Matrix.

Nach Bereinigung der Gesamtstichprobe verbleiben die Daten von 206 Jungen zur Verrechnung.

Zur Analyse der Motivierungsdaten benutzen wir die eigene Modifikation eines Itemanalyseprogramms für binär codierte Daten, des ITO1 (GMD 1976, S. 94); wir führen sie auf einer TR 440 - Anlage im Rechenzentrum der Universität des Saarlandes durch.

1.3.4.3 Ergebnisse

Die Mittelwerte der A-Statements variieren zwischen .30 und .44, die der AH-Statements zwischen .29 und .45. Die Mittelwerte der Bildsituationen variieren im Aggressionsmotiv-Teil der Testvorform zwischen .21 und .64, im Aggressions-Hemmungs-Teil zwischen .28 und .40.

Die Trennschärfen der A-Statements variieren zwischen .25 und .43, die der AH-Statements zwischen .39 und .54. Die Trennschärfen der Bildsituationen variieren im Aggressionsmotiv-Teil der Testvorform zwischen .18 und .48, im Aggressions-Hemmungs-Teil zwischen .27 und .56.

Es zeigt sich, daß von den beiden Anreizwertvariablen der aggressionsthematische Anregungsgehalt der Bildsituationen in Bezug auf Mittelwert und Trennschärfe der Statements die kritische Größe ist. Steigt er, steigen auch die Mittelwerte der A-Statements; zu den Mittelwerten der AH-Statements ist kein Zusammenhang herstellbar. Fällt er, steigt sowohl die Trennschärfe der A-Statements als auch

die Trennschärfe der AH-Statements.

Die mittlere Trennschärfe der Bildsituationen liegt für den Aggressions-Hemmungs-Teil des GITTERs mit einem Wert von .45 etwas über dem für den Aggressionsmotiv-Teil, der .37 beträgt. In beiden Teilen zeigen sich dagegen bedeutsame Trennschärfenunterschiede zwischen den Situationen, wobei die Variabilität im Aggressionsmotiv-Teil etwas größer ist.

Innerhalb der A-Statements als auch innerhalb der AH-Statements können wir keine bedeutsamen Trennschärfenunterschiede feststellen.

1.3.4.4 Diskussion

Wir halten die Ergebnisse für günstig, um mit weiteren Berechnungen fortzufahren. Speziell über die weitere Verwendung von Situationen verfügen wir bereits über gut verwertbare Hinweise. Die vorliegenden Ergebnisse weisen darauf hin, daß wir mit höheren Aussage-Trennschärfen rechnen können, wenn wir sie über einen reduzierten Satz von Bildsituationen neu bestimmen (5). Die psychometrischen Kennwerte der Aussagen beinhalten erste Hinweise auf die thematische Geschlossenheit des ihrer Entwicklung zugrunde gelegten Auswertungsschlüssels sowie auf die Güte ihrer Ableitung. Um detailliertere Aufschlüsse über sie zu erhalten, entschließen wir uns zu dem vorgesehenen nächsten Schritt, zur Untersuchung der Aussagen im faktorenanalytischen Modell.

1.3.5 Dimensionalität der konstruktthematischen Aussagen

1.3.5.1 Fragestellung

Es steht nun ein deduktiver Schritt in der Verfahrensentwicklung zur Prüfung an, die Theorieableitung der Aussagen. Bei Wahl einer dimensionsanalytischen Prüftechnik heißt das im einzelnen zweierlei, und dies ist nicht zu trennen, nämlich

- bilden die Statements eine gelungene Abbildung der den Ableitungen zugrundeliegenden Kategorien eigenen inhaltsanalytischen Gesichtspunkte und
- ist das im verwendeten Teil des A-TAT-Schlüssels zur Darstellung gebrachte Strukturmodell des Aggressionsmotiv-Systems selbst mit einer in den Motivierungsdaten auffindbaren Struktur verträglich?

Was nun den zweiten Teil der Frage anbelangt, so geht die motivationstheoretische Deutung der Aggression (vereinfacht gesagt) von der Annahme aus, daß es sich beim Aggressionsmotiv-System um ein Zweikomponentenmodell handelt, als dessen Bestandteile das Aggressionsmotiv und die Aggressions-Hemmung gelten. Es handelt sich dabei um zwei funktionell autonome, jedoch auf die gleiche thematische Kategorie (Verletzung) bezogene und somit interagierende Motive. Aufgrund qualitativ unterschiedlich getönter Lageveränderungen der damit verbundenen Affektwechsel gestaltet sich ihr Zusammenspiel zusätzlich konflikthaft im Sinne eines Annäherungs-Vermeidungs-Konflikts. Insofern die auf Annäherung gerichtete Komponente des Systems soziale Ablehnung erfährt ist es

berechtigt anzunehmen, daß die auf Vermeidung gerichtete sozial befürwortet wird. Die beiden Komponenten unterliegen somit einer ausgesprochen entgegengesetzten sozialen Wertung. Die Theorie nimmt nun weiter an, daß es sich bei den Kennwerten für die beiden Motive jeweils um eine unipolare Skala handelt und daß sie unkorreliert sind. Einander orthogonal zugeordnet, spannen die Skalen als Bezugsgrößen für individuelle Differenzen eine Fläche auf, die sich darüber hinaus zur Darstellung qualitativer Strukturbesonderheiten individueller Differenzen der Aggressivität eignet (unterschiedliche Kombinationen hoch und niedrig ausgeprägter Motivkomponenten; vgl. KORNADT 1982 a, S. 117 ff., S. 306 ff.).

Nehmen wir einmal an, die im Teil 1 der o.g. Frage zum Ausdruck gebrachte Voraussetzung sei erfüllt, so wäre es nach der Theorie plausibel, in den Motivierungsdaten eine zweidimensionale Struktur orthogonaler Faktoren zu erwarten, die im Sinne der o.g. Motive interpretiert werden könnten. Diese Erwartung wäre speziell dadurch gerechtfertigt, daß die Konstrukta Aggressionsmotiv und Aggressions-Hemmung die mit den einzelnen Kategorien jeweils zum Ausdruck gebrachten Teilaspekte einer aggressionsmotivierten Handlung auf einer hochabstrahierten, übergeordneten Ebene der Betrachtung integrativ zusammenfassen.

Es sind jedoch aus sehr viel differenzierteren Annahmen heraus auch weniger abstrakte Gesichtspunkte denkbar, die zu einer dimensionalen Untergliederung der Statements beitragen könnten,

sodaß gewissermaßen eine aufsteigend hierarchische Ordnung von Fein- bis Grob-Strukturen möglich scheint. Höherfaktorielle Strukturen werden jedoch nicht erwartet, sollen aber, sobald sie sich andeuten, ebenfalls untersucht werden.

1.3.5.2 Methode

Die Variablen der Analyse bilden die 30 Statements für das Aggressionsmotiv und die 26 Aussagen für die Aggressions-Hemmung. Zur Ermittlung der Rohwerte zählen wir aus, wie oft ein Statement über alle 11 Situationen hinweg mit JA beantwortet wird.

Wir analysieren die anfallenden Daten im faktorenanalytischen Modell (THURSTONE 1969). Die Berechnungen werden am Rechenzentrum der Universität des Saarlandes auf einer TR 440 - Anlage durchgeführt, und zwar unter Verwendung der SPSS-Prozedur FACTOR (NIE et al. 1975). Die Kommunalitäten werden iterativ geschätzt (PA2), die Rotationen erfolgen zunächst orthogonal (VARIMAX, EQUIMAX, QUARTIMAX), dann nicht-orthogonal (OBLIQUE).

1.3.5.3 Ergebnisse

Unsere eingangs formulierte Erwartung wird durch die varimaxrotierte Zweifaktorenlösung bestätigt. Diese erbringt mit 46.9 % aufgeklärter Gesamtvarianz einen Faktor "Aggressionsmotiv" (F2) sowie einen Faktor "Aggressions-Hemmung" (F1). Auf F2 entfallen 48.5 %, auf F1 51.5 % der gemeinsamen Varianz.

Die varimaxrotierte Vierfaktorenlösung stellt eine Faktorenkonfiguration dar, die ebenfalls eine theoretisch sinnvolle Interpretation gestattet. Zudem besitzt sie den höchsten Anteil an aufgeklärter Gesamtvarianz unter den orthogonalen Lösungen, die wir akzeptieren können (54.5 %). In dieser Struktur können wir - auf einem nun weniger abstrakten Niveau - innerhalb des Aggressionsmotivs eine Dimension generalisierter, erfolgszuversichtlich antizipierter Zielvorstellungen (F2) von einer Dimension allgemeiner Schemata von Handlungsmustern und handlungsbezogenen Kognitionen zur Verletzung (F3) abheben. Innerhalb der Aggressions-Hemmung gelingt uns darüber hinaus die theoretisch konzeptualisierte Differenzierung zwischen intrinsischen (F1) und extrinsischen Hemmungsanteilen (F4) im Faktorenraum. Dabei entfallen im einzelnen auf die Faktoren folgende Prozentanteile der gemeinsamen Varianz: auf F1 31.4 %, auf F2 25.4 %, auf F3 24.2 % und auf F4 schließlich 19.0 %.

Die varimaxrotierte Dreifaktorenlösung ist mit der oben dargestellten vierfaktoriellen bis auf die Aufspaltung des Aggressions-Hemmungs-Faktors identisch. Dieser zeigt sich hier in dem noch undifferenzierten Zustand, wie wir ihn aus der varimaxrotierten Zweifaktorenlösung kennen.

Die übrigen Rotationslösungen, denen bis zu sieben extrahierte Faktoren zugrunde liegen (verschiedene Abbruchkriterien, wie der Verlauf der Eigenwertdifferenzen, der Scree-Test und das Eigenwertkriterium von KAISER & DICKMAN verweisen auf bis zu sieben

bedeutsame Komponenten), klären zwar bis zu rund 59 % der Gesamtvarianz auf, stellen jedoch jeweils keine vergleichbar gut interpretierbare Gesamtkonfiguration mehr dar. Sie bestätigen die oben dargestellten Lösungsvorschläge jedoch insofern, als deren dimensionale Gliederungsgesichtspunkte (eine Ausnahme bilden die durch den Aggressionsmotiv-Split erzeugten Dimensionen) in jeweils höherfaktoriellen Lösungen erhalten bleiben, was auf die "konservative" Natur dieser Faktoren verweist. Daneben gelingt uns die Extraktion von zwei Einzelfaktoren, die geeignet sind, die varimaxrotierte Vierfaktorenlösung inhaltlich zu ergänzen. So finden wir in der fünffaktoriellen obliquen Rotationslösung (OBLIQUE, DELTA = - 2) eine Dimension allgemeiner, die intrinsischen als auch die extrinsischen Anteile der Aggressions-Hemmung umfassender Schemata von Handlungsmustern und handlungsbezogenen Kognitionen zur Aggressions-Hemmung (F1) und in der sechsfaktoriellen obliquen Rotationslösung (OBLIQUE, DELTA = - 3) eine Dimension zielerreichungsbezogener Kognitionen innerhalb des Aggressionsmotivs (F3). Damit können, wenn auch unvollkommen, zunächst einmal weitere motivationstheoretische Konstrukta dimensionsanalytisch Bestätigung finden.

Wir stellen die supplementären Gesichtspunkte, die sich aus den obliquen Analysen ergeben haben, jedoch vorläufig hintan und beschränken uns im weiteren auf den Hauptbefund dieser ersten Berechnungen. Als diesen betrachten wir die Bestätigung unseres eingangs formulierten Zweikomponenten-Modells als auch die aufgezeigte

Möglichkeit einer differenzierteren Betrachtung der Motivierungsdaten innerhalb eines vierfaktoriellen Strukturmodells.

1.3.5.4 Diskussion

Die Befunde bestätigen, daß eine vergleichsweise gelungene Theorieableitung der Aussagen vorliegt, da wir das der Entwicklung der A-TAT-Kategorien zugrunde gelegte Strukturmodell des Aggressionsmotiv-Systems mit unseren Daten replizieren konnten. Darüber hinaus haben weitere, theoretisch begründbare Gliederungsgesichtspunkte des A-TAT-Schlüssels in unseren Motivierungsdaten ein dimensionales Äquivalent gefunden.

Zusammengefaßt lassen sich die Befunde dahingehend interpretieren, daß man die Motivierungsdaten mit einer quasi unterschiedlichen "Brennweiteneinstellung" (Anzahl der extrahierten Faktoren) inspizieren und dabei auch zu theoretisch vernünftigen "Bildern" (Strukturen) kommen kann, die sich durch einen unterschiedlichen Abstraktions- und Strukturierungsgrad voneinander abheben.

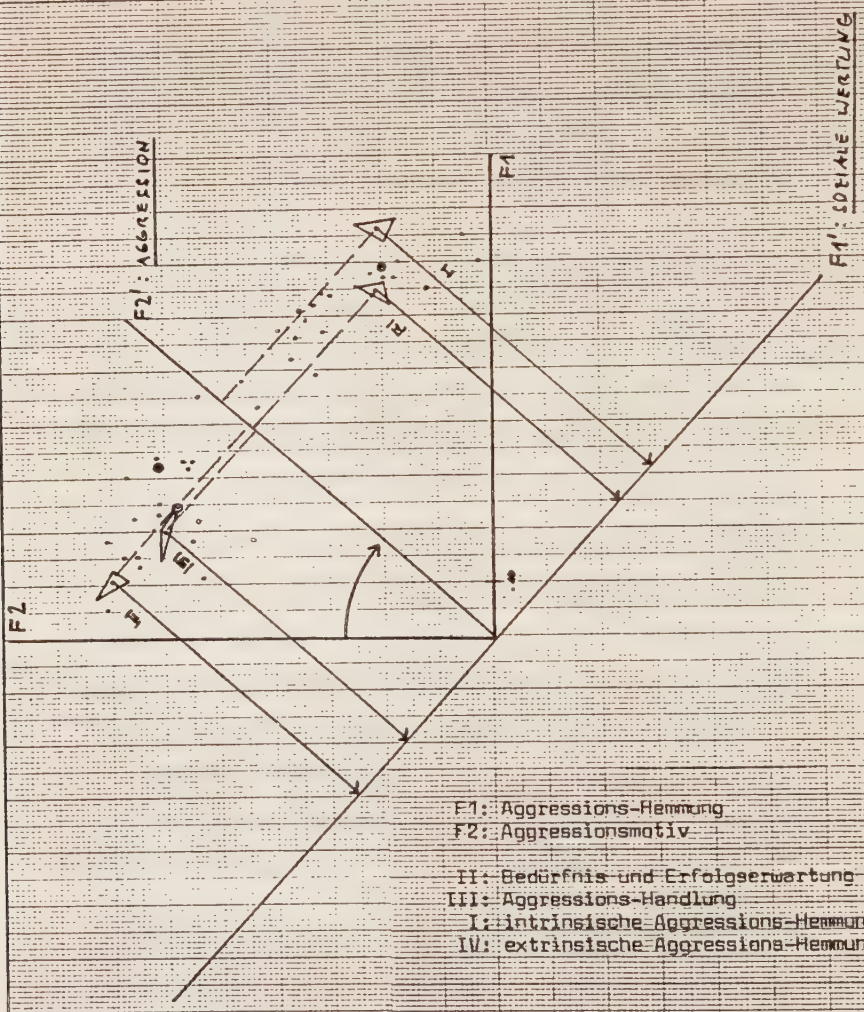
Die Frage ist, welches Betrachtungsniveau für Messungen im Aggressionsmotiv-System welche Relevanz besitzt. Wir wollen im folgenden versuchen, diese schwierige Frage einer schrittweisen Klärung zuzuführen. Konkret heißt das, eine Antwort darauf zu finden, was Motivmessung im zwei- bzw. im vierfaktoriellen Strukturmodell bedeutet als auch, durch welche

(theoretischen) Begriffe sich die Übergänge vom zwei- zum vierfaktoriellen Modellansatz beschreiben lassen. Es entspricht nicht unserer Absicht, verschiedenfaktorielle Ansätze relativ unverbunden nebeneinander existieren zu lassen.

1.3.5.5 Ein alternatives Modell

Im folgenden soll der Versuch unternommen werden, die theoretisch wesentlichen Aspekte der beiden unterschiedlich hoch dimensionierten Rotationslösungen, die wir akzeptiert haben, in einem alternativ vorgelegten Modell zusammenzufassen.

Wir gehen dazu von der varimaxrotierten Zweifaktorenlösung aus und stellen diese graphisch dar (vgl. Abb. 4). F2 bezeichnet den Faktor "Aggressionsmotiv", F1 den Faktor "Aggressions-Hemmung". Dabei fällt uns auf, daß die Endpunkte der Ladungsvektoren alle innerhalb des 1. Quadranten liegen. Sie bilden zudem ein "schmales Band", das - in beträchtlichem Abstand vom Ursprung - parallel zur 2. Winkelhalbierenden verläuft. Wir nutzen diese Information, die u.E. einen deutlichen Hinweis auf die Existenz eines Generalfaktors darstellt, wie folgt:
Wir rotieren das System F1/F2 um 41 Grad im Uhrzeigersinn graphisch nach. Danach erhalten wir einen unipolaren (F2') und einen bipolaren Faktor (F1'). F2' ist nun als ein Faktor interpretierbar, der der thematischen Kategorie des gesamten Motiv-Systems entspricht ("AGGRESSION") und F1' als eine Dimension, die die jeweils unterschiedlich gerichtete soziale Wertung der beiden Subsysteme



F1: Aggressions-Hemmung
F2: Aggressionsmotiv

II: Bedürfnis und Erfolgserwartung Aggression
III: Aggressions-Handlung
I: intrinsische Aggressions-Hemmung
IV: extrinsische Aggressions-Hemmung

Abb. 4 Graphische Nachrotation der varimaxrotierten Zweifaktorenlösung um 41 Grad UZS und integrative Aspekte des Generalfaktormodells F1'/F2. (Voruntersuchung, JUNGEN, n = 206)

verkörpert ("SOZIALE WERTUNG").

Nun kennzeichnen wir die Markiervariablen aus der vierfaktoriellen Rotationslösung (VARIMAX) und wählen dazu die jeweils drei höchsten Ladungen pro Faktor aus. Anschließend gruppieren wir die Markiervariablen zu Dreiecks-Cluster, die wir - analog den zugehörigen Faktoren - mit I bis IV kennzeichnen. Wir ermitteln nun den jeweiligen Schwerpunkt der Cluster (was nach der vektoralgebraischen Schwerpunktsformel für ein Dreieck der jeweiligen Bildung des Mittelwertevektors der Faktormarkierungsvektoren entspricht) und bilden von dort die zugehörigen Orthogonalprojektionen auf F1'.

Wir sehen nun, daß der Faktor F1' die gleiche Differenzierungsleistung erbringt wie die beiden Rotationslösungen zusammen. Zunächst trennt er die beiden Aggressionsmotiv-Faktoren (II, III) von den beiden Aggressions-Hemmungs-Faktoren (I, IV) ab. Insofern entsprechen die beiden Äste des Faktors F1' dem zwischen Aggressionsmotiv und Aggressions-Hemmung dominanten faktoriellen Differenzierungsgesichtspunkt. Darüber hinaus trennt jeder der beiden Äste innerhalb der Motivsystem-Komponenten, und zwar wie folgt: Die auf Bedürfnis und Erfolgserwartung bezogenen Anteile des Aggressionsmotivs und die intrinsischen Anteile der Aggressions-Hemmung liegen jeweils mehr nach außen, die auf instrumentelle Tätigkeiten zur Verletzung bezogenen Anteile des Aggressionsmotivs und die extrinsischen Anteile der Aggressions-Hemmung dagegen mehr zur Mitte hin. Gleichzeitig sehen wir jedoch auch, daß die analog

durchgeführten Orthogonalprojektionen auf F2' nicht zum gleichen Bild führen, obgleich wir hier feststellen können, daß die Mittelwertevektoren der Faktoren I bis IV den Faktor F2' in etwa absteigender Reihenfolge laden. Das ist jedoch nicht verwunderlich, da wir die Faktoren I bis IV sukzessiv nach der Stärke ihres Varianzbeitrages extrahiert haben. Wir werden später theoretisch zu begründen versuchen, warum der Faktor F2' (AGGRESSION) nichts zur Binnendifferenzierung im Aggressionsmotiv-System beiträgt bzw. beitragen sollte.

Damit lassen sich nun folgende Annahmen formulieren: Aufgrund einer entgegengesetzt gerichteten sozialen Wertung motivrelevanter Aspekte, die mit den A-Statements einerseits und mit den AH-Statements andererseits zum Ausdruck gebracht werden, haben sich die Faktoren "Aggressionsmotiv" und "Aggressions-Hemmung" extrahieren lassen. Aufgrund unterschiedlich hoher Ladung, die die einzelnen Konstrukt-Indikatoren auf der Dimension SOZIALE WERTUNG aufweisen, haben sich innerhalb der o.g. Systemkomponenten noch einmal weitere Faktoren verselbständigt. Unter der Voraussetzung etwa gleich hoher Ladung der thematischen Dimension AGGRESSION haben somit Unterschiede auf der Dimension SOZIALE WERTUNG die Möglichkeit geschaffen, die Korrelationsstruktur zwischen den Motivierungsdaten unter Heranziehung unterschiedlich hoch dimensionierter Modelle erklären zu können (6).

1.3.6 Probeweise Berechnung von Skalen

1.3.6.1 Wahl des Modells

Die vorliegenden Ergebnisse lassen die Wahl zwischen drei Modellen, die der Berechnung von Skalen zugrundegelegt werden können. Die zunächst durchgeführten Probeselektionen gehen vom Zweikomponentenmodell des Aggressionsmotiv-Systems aus (varimaxrotierte Zweifaktorenlösung) und sind detailliert in BURKHARDT (1978) sowie in KORNADT, ZUMKLEY & BURKHARDT (1982) dargestellt. Sie zeigen, daß auf der vorliegenden Datengrundlage zwei Subskalen für das Aggressionsmotiv und die Aggressions-Hemmung selektiert werden können, die sich jeweils durch eine hohe interne Konsistenz auszeichnen. Schwachstellen zeigen sich in einer schwachen bis mittleren wechselseitigen Korrelation sowie in den jeweiligen Rohwerteverteilungen.

Der nächste Schritt besteht nun logischerweise darin, die beiden übrigen Modelle zur Ableitung von Skalen zu prüfen. Dies soll jedoch nicht "blind" geschehen, sondern im Rahmen der Zusammenhangsannahmen, die wir zwischen den Modellen formuliert haben. Dazu gehen wir von folgenden allgemeinen Überlegungen aus:

Sollte das auf Statement-Ebene gefundene Generalfaktormodell auch auf Subskalen-Ebene die aufgezeigten integrativen Eigenschaften besitzen, so müßte es auch der Interkorrelationsstruktur von Skalen zugrunde liegen, die zum einen Teil aus dem o.g.

Zweikomponentenmodell, zum anderen aus dem darüber hinaus vorgestellten Vierkomponentenmodell des Aggressionsmotiv-Systems (varimaxrotierte Vierfaktorenlösung) heraus abgeleitet sind. Sollte dies der Fall sein, hätte das zur Folge, daß sich die Frage der Wahl (theoretisch) nicht mehr stellt.

Die Selektion der nachfolgend vorgestellten Skalen geht nun, vereinfachend gesagt, von dem Leitgedanken aus, zunächst einmal die oben angestellten Überlegungen zu prüfen. Sollten sie sich bestätigen lassen, würde das die weitere Entwicklung unseres Verfahrens vereinfachen.

1.3.6.2 Zusammenstellung der Skalen

Wir bilden nun insgesamt folgende Skalen: Auf Grundlage der varimaxrotierten Zweifaktorenlösung stellen wir eine Skala für das Aggressionsmotiv (A) und eine für die Aggressions-Hemmung (AH) zusammen. Darüber hinaus versuchen wir erstmals - von diesen Skalen ausgehend - durch weitere Ableitung Kennwerte für die theoretischen Konstrukta "Netto-Aggression" (n-A.) und "Gesamt-Aggressivität" (Ges.-A.) zu berechnen und mitzuuntersuchen (vgl. hierzu auch KORNADT 1982 b, S. 27 f.). Auf Grundlage der varimaxrotierten Vierfaktorenlösung bilden wir vier Skalen, die jeweils einen der Faktoren schwerpunktmäßig laden sollen. Wir bezeichnen sie der Einfachheit halber nach den Faktoren mit F1 bis F4.

Es bestent nun eine besondere Beziehung zwischen den beiden o.g.

Strukturmodellen dergestalt, daß sich die in der vierfaktoriellen Rotationslösung aufgezeigten Dimensionen gewissermaßen als "splits" der in der zweifaktoriellen bestehenden auffassen lassen. Wir versuchen diese Besonderheit auf Subskalen-Ebene so zu berücksichtigen, daß sich durch Halbierung der im zweifaktoriellen Modell definierten Skalen (A, AH) die jeweils zugehörigen "Split-Skalen" (F_i) erzeugen lassen.

Das hat nun insgesamt eine einfache algebraische Beziehung zwischen den Skalen zur Folge; und zwar ist:

A	=	F2	+	F3
AH	=	F1	+	F4
n-A.	=	A	-	AH
Ges.-A.	=	A	+	AH

Auf die Skalen F1 bis F4 entfallen demzufolge folgende Statements:

F1 (intrinsische Aggressions-Hemmung):

- Er möchte sich wegen des unerfreulichen Vorfalls entschuldigen.
- Er denkt: "Der meint es garnicht so!", und man einigt sich in Güte.
- Um eine Auseinandersetzung zu verhindern, entschuldigt er sich.
- Sie haben ein schlechtes Gewissen und hören auf, zu streiten.

F2 (Bedürfnis und Erfolgserwartung Aggression)

- Er möchte mit ihm streiten.
- Er möchte ihm mal die Meinung sagen.
- Er denkt: "Dem werd' ich's geben!"
- Er möchte ihn verprügeln.

F3 (Aggressions-Handlung):

- Sie schimpfen miteinander.
- Es kommt zu einem Streit.
- Sie verhauen sich.
- Sie geraten aneinander.

F4 (extrinsische Aggressions-Hemmung):

- Er möchte die Folgen eines Streits vermeiden.
- Er denkt: "Hoffentlich kommt es zu keinem Streit, denn mit dem ist nicht zu spaßen!"
- Um unliebsame Folgen zu vermeiden, läßt er sich nicht auf einen Streit mit ihm ein.
- Er möchte sich nicht mit ihm einlassen, um ernsthafte Folgen zu verhindern.

Vor allem zur Verbesserung der Rohwerteverteilungen wählen wir dazu einheitlich für alle Skalen die Bildsituationen mit den Nummern 1, 2, 5, 9, 10 und 11.

1.3.6.3 Schätzung einiger Skalenparameter

Wir ermitteln den Rohwert der Skala F1, indem wir auszählen, wie oft die zugehörigen Statements über alle 6 Bildsituationen hinweg bejaht werden. Ebenso ermitteln wir die Rohwerte auf den übrigen F-Skalen. Die Rohwerte auf den restlichen Skalen lassen sich wie dargestellt rechnerisch ableiten.

Diese Rohwerte ziehen wir zunächst zur Schätzung einiger wichtiger Skalenparameter heran (vgl. Tab. 2).

SKALA	$\bar{x}$	s	r_{tt}
F1	8.8	5.5	.83
F2	9.9	4.5	.81
F3	9.0	4.4	.82
F4	10.5	5.3	.82
A	18.8	7.6	.87
AH	19.3	9.2	.87
n-A.	-0.5	9.5	.79
Ges.-A.	38.0	14.0	.91

$\bar{x}$: arithmetisches Mittel
s : Standardabweichung
r_{tt}: interne Konsistenz (split-half, aufgewertet nach SPEARMAN-BROWN)

Tab. 2 Psychometrische Kennwerte einiger probeweise selektierter Subskalen der Testvorform des AM-GITTERs (JUNGEN; n = 206; TR/IT01)

Die in Tabelle 2 aufgeführten Reliabilitäten für die Skalen n-A. und Ges.-A. werden als Verhältnis der Varianz der wahren zur Varianz der beobachteten Werte nach den Regeln zur Berechnung linear abgeleiteter Testwerte (vgl. TACK 1973, S. 21 ff.) bestimmt. Wir setzen dabei für die Reliabilitäten der Skalen A und AH die oben aufgeführten r_{tt}-Werte ein und bestimmen deren Korrelation als Rangkorrelation (rho; vgl. Tab. 3).

1.3.6.4 Dimensionalität der Skalen

Wir ermitteln nun die Rangkorrelationen zwischen den definierten Skalen (vgl. Tab. 3) und untersuchen die resultierende rho-Matrix im faktorenanalytischen Modell.

rho	F1	F2	F3	F4	A	AH	n-A.	Ges.-A.
F1		.27	.30	.45	.32	.86	-.54	.73
F2			.48	.39	.85	.39	.28	.69
F3				.18	.85	.28	.42	.63
F4					.30	.83	-.54	.72
A						.37	.42	.76
AH							-.64	.86
n-A.								-.21
Ges.-A.								

Tab. 3 Korrelationsmatrix der Subskalen der Testvorform (JUNGEN; n = 206; SI/SPSS, NONPAR CORR)

Eine spezielle Charakteristik der Eigenwertverlaufsfunktion (Scree-Test; vgl. ÜBERLA 1977, S. 127 ff.) verweist deutlich auf zwei Faktoren, die extrahiert werden sollten, wie Abbildung 5 verdeutlicht.

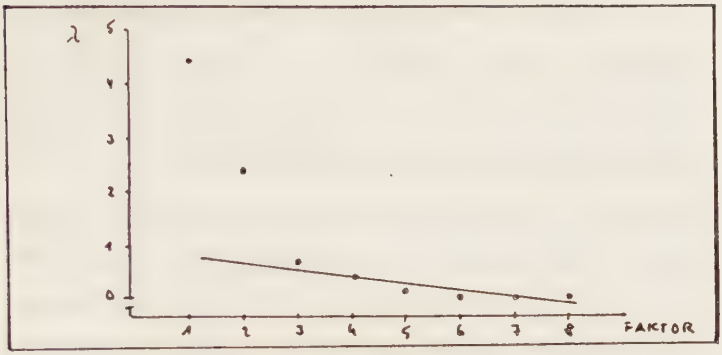


Abb. 5 Eigenwerte der rho-Matrix, der Größe nach geordnet und SCREE-TEST

Wir extrahieren darauf zwei Faktoren und erhalten nach einer Iteration folgende Ladungsmatrix (vgl. Tab. 4):

	UNROTIERTE FAKTOREN		
VARIABLE	Faktor 1	Faktor 2	h^2
F1	.75	-.39	.72
F2	.70	.47	.71
F3	.62	.60	.74
F4	.74	-.39	.70
A	.76	.65	.99
AH	.88	-.46	.99
n-A.	-.23	.93	.92
Ges.-A.	.98	.02	.97
Summe a_j^2	4.36	2.39	
% gemeinsame VARIANZ :	64.60	35.40	
% Gesamt-VARIANZ :	54.50	29.88	SUMME: 84.38

Tab. 4 Ladungsmatrix der Subskalen der Testvorform (JUNGEN; n = 206; SI/SPSS, FACTOR (PA2))

Wie wir Tabelle 4 entnehmen können, läßt sich unsere eingangs geäußerte Vermutung bestätigen. Der unrotierten Ladungsmatrix liegt eindeutig ein Generalfaktormodell zugrunde, das die gesamte Subtestvarianz zu einem hohen Anteil (84.38 %) erklärt.

Bei Faktor 1 handelt es sich um einen fast doppelt so varianzstarken Faktor wie bei Faktor 2, und wir können ihn als thematischen Faktor AGGRESSION identifizieren. Motivationstheoretisch betrachtet gewährleistet eine (bedeutsame) Ladung auf dieser Dimension die Motivspezifität einer (fundamentalen) Subskala.

In dem bipolaren Faktor 2 erblicken wir die Dimension der SOZIALEN WERTUNG unterschiedlich gerichteter Motivkomponenten des Aggressionsmotiv-Systems. Subskalen für Aggressions-Hemmung, die sozial befürwortet wird, korrelieren demzufolge mit dieser Dimension unter einem anderen Vorzeichen als Subskalen für das Aggressionsmotiv, das eine soziale Ablehnung erfährt. Motivationstheoretisch gesehen deuten Unterschiede in der Ladungsrichtung auf dieser Dimension (positive vs. negative Korrelation mit dieser Dimension) bei Subskalen darauf hin, daß die durch sie erfaßten Konstrukta einer sozial entgegengesetzten Wertung unterliegen.

Mit der Bestimmung der Dimensionalität dieser probeweise selektierten Subskalen des AM-GITTERs schließen wir die Voruntersuchungen ab.

1.4 Zusammenfassung

In einer Reihe erster Schritte zur Entwicklung des AM-GITTERs haben wir zunächst Überlegungen dazu angestellt, wie wir die situationsseitigen einerseits und die personseitigen Bestandteile des Gitters andererseits darstellen könnten. Wir haben uns daraufhin entschlossen, zur Darstellung der situationsseitigen Bestandteile die Bilder eines Aggressions-TAT für Jungen heranzuziehen und die personseitigen Bestandteile aus dem Auswertungsschlüssel eines Aggressions-TAT für männliche Jugendliche und Erwachsene abzuleiten. Zur Darstellung letzterer wählten wir die Form kleiner Sätze, die Konstruktcharakter zum Teil für das Aggressionsmotiv, zum Teil für die Aggressions-Hemmung besitzen und dem Probanden eine individuelle, motivthematische Strukturierung der o.g. Bilder bzw. Bildsituationen ermöglichen sollten.

Die Prüfung der Bilder erfolgte in zwei Abschnitten. Wir haben sie zunächst TAT-analog vorgelegt und dazu Geschichten erzählen lassen. Die Auswertung der Geschichten ergab, daß die Bilder zur Anregung von Phantasieinhalten geeignet sind, die sich auf feindselige Aggression zwischen gleichaltrigen Jungen und auf deren Hemmung beziehen. Fast durchgehend bieten die Bilder alternativ dazu Anhaltspunkte an, mit denen eine Art freundschaftsbezogener Motivierung verknüpft wird. Der nächste Schritt bestand in der Prüfung der Bilder als Bestandteile einer Testvorform des AM-GITTERs. Dabei zeigte sich, daß sie als situationsseitige Bestandteile

von Aggressionsmotiv- und Aggressions-Hemmungs-Skalen unterschiedlich gut geeignet sind. Unter psychometrischen Gesichtspunkten erwies sich ein mittlerer Anregungsgehalt für Aggression (-smotivthematik) als vorteilhaft.

Mittels eines Experten-Rating versuchten wir, aus einer umfangreichen Liste kleiner Sätze diejenigen herauszufinden, die unter differenzierten theoretischen Gesichtspunkten den vorgesehenen Konstruktcharakter am besten zum Ausdruck brachten. Anschließend untersuchten wir die Sätze ebenfalls als Bestandteile einer Testvorform. Dimensionsanalytische Untersuchungen zeigten, daß wir mit unseren Sätzen wichtige theoretische Gesichtspunkte der internen Struktur des A-TAT-Schlüssels, aus dem wir sie abgeleitet hatten, zum Ausdruck gebracht hatten. Zunächst konnten wir mit den personseitigen Bestandteilen der Testvorform die Zweikomponenten-Theorie des Aggressionsmotiv-Systems sensu KORNADT bestätigen. Darüber hinaus konnten wir ein differenzierteres Vierkomponentenmodell aufzeigen und anschließend beide Strukturen in einem alternativ vorgeschlagenen Generalfaktormodell des Aggressionsmotiv-Systems vereinigen.

Proberechnungen zur Bildung von Motivskalen in der Testvorform verwiesen auf Möglichkeiten ein Testverfahren zu erstellen, das nicht nur mit motivationstheoretischen, sondern auch mit psychometrischen Richtlinien in Einklang gebracht werden konnte. Die dimensionsanalytische Prüfung eines probeweise zusammengestellten AM-GITTERs bestätigte das o.g. Generalfaktormodell des Aggressionsmotiv-Systems.

2. Hauptuntersuchung und Konstruktion des AM-GITTERs

2.1 Einleitende Gesichtspunkte

Die Voruntersuchungen sollten zunächst einmal klären, ob die von uns entworfene Strategie der spezifischen Gestaltung eines Aggressions-Motiv-Gitters unter streng psychometrischen Gesichtspunkten betrachtet einen gangbaren Weg darstellt. Darüber hinaus sollten sie Datenmaterial bereitstellen, um spezifische Fragen konstruktiver Schritte im Detail aufzuspüren und zu beantworten.

Man kann die in diesem Zusammenhang eruierten Befunde summarisch dahingehend zusammenfassen, daß sie eine Bestätigung für den von uns eingeschlagenen Weg darstellen. Dabei war jedoch auch nach den nun vorliegenden Ergebnissen noch offen, ob und in welchem Umfang wir im weiteren Revisionen der Bestandteile unserer ersten Testvorform (Bildsituationen, Aussagen) in Betracht ziehen mußten.

Wir entschieden uns an dieser Stelle aufgrund der vorliegenden Daten und Ergebnisse, die insbesondere auch im Hinblick auf die recht hohen Leistungsanforderungen an die Probanden bei der Bearbeitung des sehr umfangreichen ersten Itempools zu werten waren, vorerst ohne größere Veränderungen in den konstruktiven Bestandteilen unserer Testvorform fortzufahren, d.h. im wesentlichen ein erheblich verkürztes Verfahren erneut vorzulegen und zu prüfen.

Hinsichtlich der Dimensionalität der mit der Testvorform erhobenen Motivierungsdaten hatten sich Perspektiven gezeigt, die in dieser Form zunächst nicht erwartet worden waren. Da wir deren Klärung grundlegende Bedeutsamkeit für die theoretische Begründung des Verfahrens einräumten, mußten die weiteren Schritte diesbezüglich sehr sorgsam eingeleitet werden. Wesentlich dabei schien uns die Beantwortung der Frage, als wie sicher (invariant) die auf der Grundlage einer erstmaligen größeren Datenaufnahme beruhenden Aussagen einzuschätzen waren. Erst danach konnte entschieden werden, nach welchem Modell (bzw. nach welchen Modellen) Skalen entwickelt werden sollten und deren internes Beziehungsgeflecht (Interkorrelationsstruktur) zu beschreiben war.

2.2 Zusammenstellung einer neuen Testvorform

Es stellt sich nun die Frage, nach welchen Überlegungen im einzelnen eine neue, reduzierte Testvorform zusammengestellt und an einer neuen, umfassenderen Stichprobe überprüft werden soll. Die in diesem Zusammenhang angestellten Überlegungen und initiierten Schritte lassen sich in zwei Teilschritte untergliedern:

(a) Unter selektivem Aspekt scheiden wir zur weiteren Untersuchung die Bildsituationen und Aussagen aus, die sich gegenüber den anderen weniger gut bewährt haben.

(b) Unter konstruktivem Aspekt führen wir Änderungen in den neutralen Aussagen und im Test-Layout durch.

Näheres dazu ist den folgenden Teilabschnitten zu entnehmen.

2.2.1 Selektive Maßnahmen

Den neuen Satz von Bildsituationen stellen wir unter dem Gesichtspunkt zusammen, für die grundlegenden Motivskalen A (Aggressionsmotiv) und AH (Aggressions-Hemmung) möglichst trennscharfe und normalverteilte Meßwerte zu erhalten. Speziell hegen wir dabei die Zielvorstellung, mit beiden Skalen im unteren Bereich der Merkmalsausprägung besser differenzieren zu können, als bisher (vgl. KORNADT, ZUMKLEY & BURKHARDT 1982). Zu diesem Zweck scheint uns die Bildsituationsgruppe, bestehend aus den Nummern 1, 2, 5, 9, 10 und 11 am geeignetsten, zumal sie sich auch zur Berechnung unserer "Split-Skalen" (F1 bis F4) in den Voruntersuchungen gut bewährt hat.

Die Aussagen für die neue Testvorform werden schwerpunktmäßig so ausgewählt, daß wir bei der faktoriellen Analyse der neuen Motivierungsdaten auf Statement-Ebene sowohl im Hinblick auf eine zweifaktorielle als auch auf eine vierfaktorielle orthogonale Rotationslösung bestmögliche und theoriekonsistente dimensionale Einfachstruktur der Statement-Ladungsvektoren und damit der gesamten Ladungsmatrix erwarten können.

Mit der Replizierbarkeit der in der Voruntersuchung vorgestellten zweifaktoriellen orthogonalen Rotationslösung wird implizit auch das Generalfaktormodell bestätigt. Über dessen Bestätigung auf Subskalen-Ebene werden wir erneut Berechnungen anzustellen haben, analog den schon im Rahmen der Voruntersuchungen durchgeführten.

Zusammengefaßt: Differenzierungsfähigkeit und Meßgenauigkeit der neuen Skalen versuchen wir primär über Manipulation der Bildsituationen, ihre Dimensionalität dagegen durch Manipulation der Aussagen zu steuern.

2.2.2 Konstruktive Maßnahmen

Die ursprünglich "neutralen" Puffer-Statements sollen in der neuen Testvorform in die Dimensionierungsanalysen miteinbezogen werden und dabei aus Gründen der internen Validität der Befunde einen eigenen Faktor aufspannen. Dazu wird ein Teil der Puffer-Statements ausgeschieden, ein weiterer in der bisherigen Form beibehalten und einige werden gänzlich umformuliert. Insgesamt werden sie dadurch - einigermaßen gezielt - auf den thematischen Bereich "FREUNDSCHAFT" abgestimmt, weil

(a) die im Rahmen der Voruntersuchungen durchgeführten Inhaltsanalysen der zu den LESSERschen Bilder produzierten TAT-Geschichten gezeigt hatten, daß die Bildsituationen neben AGGRESSION, sozusagen als alternatives Thema auch FREUNDSCHAFT aktivieren können und

(b) aufgrund dieser Besonderheit zudem die Aussicht bestand, die diagnostische Bandbreite des AM-GITTERs (in diesem Situationsbereich) eines Tages evtl. um ein nichtaggressionsthematisches Motiv, das womöglich als eine Alternative zur Aggressivität im Sinne einer (inkompatiblen) prosozialen Motivation aufzufassen wäre (vgl. KORNADT 1981 b), erweitern zu können.

Im Hinblick auf das Test-Layout vereinigen wir nun Bilder und

Aussagen in einem gemeinsamen Testheft dergestalt, daß auf einer DIN A 4 - Seite im Querformat jeweils links ein Bild (Breite: 9.5 cm; Höhe: 15 cm) und rechts daneben die Aussagen aufgeführt sind. Daneben entwerfen wir ein Deckblatt im gleichen Format, auf dessen linker Seite die üblichen bio- und soziographischen Daten eines Probanden eingetragen werden können und auf dessen rechter Seite der Text zur Einführung in die Bearbeitung des AM-GITTERs (vgl. Abschnitt 7.1 Teil I dieser Arbeit) zu lesen ist. Zuletzt kreieren wir ein Codierblatt für die Gesamtauswertung des Testverfahrens und stellen es in den Anhang des Testheftes. Zusätzlich sehen wir pro Bildsituation - unterhalb der Aussagen - Möglichkeiten zur Zwischencodierung für das Aggressionsmotiv (A), für die Aggressions-Hemmung (AH) sowie für die einzelnen Faktoren der in der Voruntersuchung extrahierten vierfaktoriellen Rotationslösung (F1 bis F4) vor.

Das auf diese Art und Weise zusammengestellte AM-GITTER (2. Testvorform) besteht nunmehr im einzelnen aus

(a) 6 Bildsituationen (mit der Nummernfolge 1-2-5-9-10-11 im Test)

(b) 35 Aussagen, davon 14 für das Aggressionsmotiv (7+7 für d. splits)
14 für d. Aggressions-Hemmung (----- * -------)
7 für den thematischen Bereich FREUNDSCHAFT

(c) einem Deck- und einem Codierblatt.

Die unter (b) dargestellte personseitige Aufteilung des AM-GITTERs ist nicht zuletzt auch im Hinblick auf Überlegungen von BERKOWITZ (1962; zit. n. SELG 1974, S. 61) sinnvoll, der Schwierigkeiten bei

der Messung der Aggressivität dadurch erwartet, daß Schuldgefühle und aggressionsneutrale, durch das Testmaterial ausgelöste Reaktionstendenzen intervenieren und den Ausdruck von Aggression im Test hemmen; sie erlaubt dem Probanden nämlich eine im Rahmen der aufgefundenen Möglichkeiten vergleichsweise umfassende thematische Auffassung der Bildsituation(en) und läßt ihm die Wahl, diese auch differenziert zum Ausdruck bringen zu können (durch JA-Antworten), und nicht etwa (hemmungsbedingt oder durch sonstige Prozesse veranlaßt) zurückhalten zu müssen (durchgehend NEIN-Antworten bei nicht-aggressionsmotivthematischer Strukturierung).

Zu den Bildsituationen vgl. ANHANG I, zu den konstruktthematischen und neutralen Aussagen vgl. ANHANG II.

2.3 Durchführung der Untersuchung

2.3.1 Stichprobe

In fünf kleinen bis mittelgroßen Städten des Saarlandes wählen wir 12 Grundschulen zur Zusammenstellung einer Stichprobe aus. Nach Abklärung der formaljuristischen Auflagen beziehen wir alle dort eingerichteten 4. Klassen, insgesamt 31, in die Untersuchung ein.

Da wir das Verfahren jeweils im Gruppenversuch vorzulegen beabsichtigen, scheiden wir nach Rücksprache mit den Klassenlehrern

diejenigen Kinder von der Untersuchung aus, bei denen nach deren Einschätzung ein Verständnis des Inhaltes beim Lesen einfacher Texte nicht hinreichend gewährleistet werden kann. Es handelt sich dabei um Kinder, deren Muttersprache nicht Deutsch ist sowie um einige, die Förderkurse in Lesen und Rechtschreiben besuchen. "Ausländer"-Kinder, die nach Meinung der Klassenlehrer dem o.g. Kriterium entsprechend in die Stichprobe einbezogen werden könnten, werden grundsätzlich mituntersucht, jedoch bezüglich einer evtl. erforderlichen nachträglichen Bereinigung der Stichprobe als solche gekennzeichnet. Danach verbleiben für die Endauswertung der Daten 282 Jungen.

Wir sondern diesmal nicht mehr, wie in den Voruntersuchungen, die Mädchen einer Klasse aus, sondern legen diesen ebenfalls - aus rein explorativen Gründen - ein AM-GITTER vor. Sie werden zuvor nach der gleichen Methode (s.o.) ausgelesen; es verbleiben danach noch 284 Mädchen.

Die Altersverteilung der Kinder zum Zeitpunkt der Untersuchung ist der Tabelle 5 zu entnehmen.

ALTER in Jahren	8	9	10	11	12	SUMME
JUNGEN (Anzahl)	5	183	82	11	1	282
MÄDCHEN "	3	212	57	12	0	284

Tab. 5 Altersverteilung der Kinder in der Stichprobe zur Hauptuntersuchung

2.3.2 Datenaufnahme

Die Datenaufnahme erfolgt im Zeitraum September bis November 1980. Die Testungen werden morgens während der offiziellen Unterrichtszeit in eigens dazu bereitgestellten Unterrichtsräumen und im Gruppenversuch durchgeführt. Dazu stehen zwei Unterrichtsstunden zur Verfügung (i.d. Regel die bd. ersten). Die Klassen einer Schule werden nach Möglichkeit jeweils am gleichen Tag getestet und in einer Gruppe zusammengefaßt. Alle Testungen werden vom gleichen Testleiter durchgeführt. Lehrpersonen sind nicht anwesend, besondere Vorkommnisse nicht zu vermerken.

Die Instruktion zur Bearbeitung des Verfahrens erfolgt standardmäßig (vgl. Abschnitt 7.1, Teil I). Am Schluß werden alle Testhefte auf formale Bearbeitungsfehler kontrolliert und verbessert.

Das Verfahren wird in zwei Versionen (A, B) vorgelegt. Version B unterscheidet sich von Version A dadurch, daß die Bildsituationen (mit unveränderter, über alle Bilder gleichbleibender Statement-Abfolge) in umgekehrter Reihenfolge vorgegeben werden. Durch diese Vorkehrung soll ein gegebenenfalls auftretender Ermüdungseffekt ausbalanciert werden. Ferner erhalten nebeneinandersitzende Kinder stets unterschiedliche Versionen zur Bearbeitung.

2.4 Prüfung der Bildsituationen und Aussagen

2.4.1 Fragestellung

Wir wollen die Bildsituationen und Aussagen der zweiten Testvorform zunächst wieder durch einfache psychometrische Kennwerte (Mittelwert, Trennschärfe) charakterisieren. Die Kennzeichnung der Bildsituationen soll für die Motivbereiche Aggression und Aggressions-Hemmung getrennt vorgenommen werden.

2.4.2 Methode

Die wenigen noch verbliebenen "missing values" werden alternierend mit 0/1 gescort. Die Mittelwerte und Trennschärfen errechnen wir nach der in der Voruntersuchung erprobten Art und Weise (vgl. Abschnitt 1.3.4.2).

2.4.3 Ergebnisse

Die Ergebnisse sind im einzelnen den Tabellen 6 und 7 zu entnehmen.

SITUATION Nr.	AGGRESSIONSMOTIV p	AGGRESSIONSMOTIV r_{bis}	AGGRESSIONS-HEMMUNG p	AGGRESSIONS-HEMMUNG r_{bis}
1	.25	.44	.54	.44
2	.60	.38	.37	.44
5	.54	.43	.44	.39
9	.37	.48	.43	.49
10	.43	.50	.46	.46
11	.24	.42	.53	.44

Tab. 6 Psychometrische Kennwerte der Bildsituationen (JUNGEN; n = 282; TR/ITO1)

AGGRESSIONSMOTIV-Statements			AGGRESSIONS-HEMMUNGS-Statements		
Auss. Nr.	p	r_{bis}	Auss. Nr.	p	r_{bis}
01	.42	.41	15	.42	.40
02	.38	.47	16	.46	.46
03	.50	.43	17	.45	.47
04	.51	.37	18	.45	.51
05	.42	.50	19	.43	.52
06	.42	.47	20	.49	.46
07	.42	.50	21	.39	.37
08	.38	.44	22	.48	.44
09	.36	.46	23	.46	.33
10	.43	.49	24	.48	.45
11	.29	.46	25	.53	.50
12	.37	.36	26	.50	.36
13	.40	.51	27	.45	.41
14	.38	.33	28	.50	.52

Tab. 7 Psychometrische Kennwerte der Aussagen (JUNGEN; n = 282; TR/IT01)

2.4.4 Diskussion

Die Ergebnisse sind insgesamt als günstig zu betrachten und bestätigen die von uns getroffene Auswahl von Bildsituationen und Aussagen. Über die Zusammenstellung endgültiger Motivskalen soll jedoch erst entschieden werden, wenn die Befunde aus den nachfolgend noch darzustellenden Dimensionierungsanalysen vorliegen. Zudem sollen dann noch mehrere Probeselektionen durchgerechnet werden, deren Ergebnisse wir dann unter gleichzeitiger Beachtung aller relevanter Kriterien miteinander vergleichen wollen, um die bestmögliche Kombination herauszufinden.

2.5 Dimensionalität der Aussagen

2.5.1 Vorüberlegungen

Wie bereits aus den entsprechenden Teilabschnitten ersichtlich wurde, besitzt die Prüfung der Dimensionalität der Aussagen, die ja den personseitigen Bestandteil des Meßinstruments vergegenständlichen, erhebliche theoretische Relevanz.

Als Konstruktindikatoren unmittelbar aus der Theorie über das zu messende Persönlichkeitsmerkmal abgeleitet, nimmt man von ihnen an, daß sie letzteres quasi punktuell an verschiedenen Stellen, an denen es in einem (motivierten) Handlungsablauf sichtbar wird, erfassen. Somit gestatten Beziehungsaspekte, die zwischen den Aussagen aufgedeckt werden, ihrerseits wieder Rückschlüsse zur Triftigkeit bestimmter Differenzierungen und Annahmen, die mit ihnen zum Ausdruck gebracht werden sollen.

Die Bündelung einzelner Indikatoren zu Indikator-Gruppen und auch die weitere Bündelung dieser Gruppen im Rahmen der Subskalen-Bildung erfolgt ebenfalls unter theoretischen Annahmen. Deswegen gestatten Beziehungsaspekte, die zwischen diesen Bündeln (Summenvektoren) aufgedeckt werden, ihrerseits wieder Rückschlüsse zur Triftigkeit der Bündelung im Rahmen der Subskalen-Bildung. So etwa ist das auch, zumindest was den ersten Teil des oben Gesagten betrifft, in der Entwicklung des Leistungsmotiv-Gitters (LM-GITTER; SCHMALT 1973, 1974, 1976 a, b) verstanden worden.

Damit wird nun deutlich, daß mit nachgewiesener Kongruenz zwischen theoretischen Annahmen zur Struktur des Aggressionsmotiv-Systems und Merkmalen empirisch in den Motivierungsdaten aufzeigbarer Strukturgefüge die Güte der Theorieableitung des Verfahrens unmittelbar faßbar wird. Dies ist ein unmittelbar einleuchtender Aspekt der (Konstrukt-) Validität des AM-GITTERs, der an dieser Stelle in dessen Entwicklung einbezogen werden soll.

Die im Rahmen unserer Voruntersuchungen durchgeführten Analysen haben den o.g. Denkansatz überraschend gut für das Aggressionsmotiv-System und die dafür entwickelte Testvorform bestätigt, und zwar sowohl für dessen Binnendifferenzierung, als auch für dessen Ableitung aus einer allgemeinen Theorie motivierten Handelns (d.h. deren Anwendung auf den thematischen Bereich Aggression), wie speziell unser alternativ vorgestelltes Generalfaktormodell zeigt. Beide Komponenten den Modells (AGGRESSION, SOZIALE WERTUNG) haben allgemeintheoretischen Charakter und sind gerade deshalb geeignet, motivsystemspezifische Begriffe bis auf die Ebene der beiden Partialkonstrukte "Aggressionsmotiv" und "Aggressions-Hemmung" zu beschreiben bzw. voneinander abzuheben.

Die nachfolgend anzustellenden Berechnungen werden somit zunächst einmal unter dem Aspekt einer empirischen Bestätigung der dem Verfahren zugrundegelegten Theorie als auch unter dem Gesichtspunkt der Güte der Theorieableitung des Verfahrens zu betrachten sein.

Da wir nun zur Formalisierung der internen Struktur der Motivierungsdaten gleich mehrere Modellvarianten aufgefunden haben, stellt sich im weiteren die Frage, ob sich diese Varianten im einzelnen erneut bestätigen (Invarianzeigenschaften der Varianten) und, wenn ja, darüber hinaus noch einmal in einer von ihnen integrativ zusammenfassen lassen, wie erstmals gezeigt. Bei positiver Befundlage würde die in diesem Zusammenhang geäußerte Verselbständigungshypothese für die einzelnen Faktoren zumindest indirekt mitbestätigt (vgl. Abschnitt 1.3.5.5 Teil II).

Von nicht zu unterschätzender Bedeutung wäre es allerdings auch zu erfahren, ob sich eine Möglichkeit bietet, die Dimensionalität der Statements als auch der Subskalen im gleichen Modell zu beschreiben, was sich ja im Generalfaktormodell als durchführbar erwies. Auf letzteres bezogen würde das heißen, daß man bei der Subskalen-Bildung in der gleichen Ebene bleibt, in der bereits die einzelnen Konstrukt-Indikatoren "liegen". Damit wäre die Interpretation der Subskalen wesentlich vereinfacht, zumal die zugehörigen Subtestvektoren ja nicht aus dem für die einzelnen Konstrukt-Indikatoren gültigen Modell "hinausragen" (wie z.B. die entsprechenden "Netto-Kennwerte" im LM-GITTER oder im A-TAT). Zudem ergäben sich daraus besondere Gesichtspunkte für die theoretische Vorgabe der Dimensionalität einzelner Skalen sowie für weiterführende Maßnahmen der Validierung.

Sollte sich unser Testverfahren wie oben vorgezeichnet formal begründen lassen, so wäre es weiter von besonderem Interesse,

die Gültigkeit des dazu herangezogenen Modells dadurch weiter aufwerten zu können, daß man Belege für seinen Allgemeinheitscharakter erbringt. Eine Möglichkeit sehen wir darin, dessen Leistungsfähigkeit an (zumindest) einem kritischen Kennwert zu überprüfen. Wir haben diesbezüglich vor, die auf den thematischen Bereich FREUNDSCHAFT abgestimmten Aussagen (F-STATEMENTS) mitzuuntersuchen, wobei wir über deren Vektorkoordinaten zuvor motivationstheoretische Annahmen formulieren. Eine weitere Möglichkeit, die Gültigkeit des Modells zu prüfen, besteht u.E. darin, daß man die Population, innerhalb derer man sich bislang bewegt hat, um eine neue erweitert, in der das Modell als ebenfalls gültig angenommen werden kann, und sodann feststellt, ob dies zutrifft. Wir wollen die auf die gleiche Weise (und unter den gleichen Bedingungen) erhaltenen Motivierungsdaten der Mädchen daher nach der gleichen Analysetechnik auswerten; dies allerdings mehr explorativ, da das Verfahren für diese Population nicht gedacht ist (7). Es braucht natürlich nicht besonders hervorgehoben zu werden, daß damit die durch das Modell beschriebenen Größen an inhaltlicher Evidenz gewinnen, sollte es auch in dieser Population Bestätigung finden.

Damit besteht eine Vielzahl gleichermaßen interessanter Einzelaspekte, unter denen die Dimensionalität nun aller Aussagen des GITTERs erneut zu prüfen ist. Von den Ergebnissen erwarten wir Aufschlüsse darüber, wie die weiteren konstruktiven Schritte sinnvollerweise eingeleitet werden können.

2.5.2 Ausgangshypothesen

Unsere im Anschluß an die Ergebnisse der Voruntersuchungen eingeleiteten konstruktiven Schritte zur Weiterentwicklung des Verfahrens (vgl. Abschnitt 2.2.1 Teil II) hatten bereits zum Ziel, das für uns günstige Gesamtresultat aus den Dimensionierungsanalysen im Retest bestätigen zu können. Demgemäß erwarten wir nun im einzelen

auf STATEMENT-Ebene:

(a) die Replizierbarkeit des mit der zweifaktoriellen orthogonalen Rotationslösung verknüpften Modells;

(b) die Replizierbarkeit des mit der vierfaktoriellen orthogonalen Rotationslösung verknüpften Modells;

(c) die Integrierbarkeit der beiden o.g. Modellvarianten in einem zweifaktoriellen (unrotierten) Generalfaktormodell;

und auf SUBSKALEN-Ebene:

(d) eine Bestätigung für das zweifaktorielle (unrotierte) Generalfaktormodell.

Für die Statements aus dem Bereich FREUNDSCHAFT erwarten wir, daß sie in (orthogonalen) Rotationslösungen einen eigenen Faktor aufspannen. Hinsichtlich der Ladungsvektoren der F-Statements im zweifaktoriellen Generalfaktormodell erwarten wir substantielle Ladungen auf dem Ast der Dimension SOZIALE WERTUNG, den auch die Aggressions-Hemmung (und vor allem deren intrinsischen Anteile) lädt; keine Ladungen jedoch auf dem Ast, den das Aggressionsmotiv lädt und darüber hinaus keine substantiellen Ladungen auf dem thematischen Faktor AGGRESSION.

Für die den Mädchen-Daten zugrundeliegenden Strukturen formulieren wir keine speziellen Erwartungen.

2.5.3 Methode

Die Variablen der Analyse bilden die 14 Aggressionsmotiv-Statements, die 14 Aggressions-Hemmungs-Statements und die 7 Freundschafts-Statements. Zur Ermittlung der Rohwerte zählen wir aus, wie oft ein Statement über alle 6 Bildsituationen hinweg mit JA beantwortet wird.

Wir analysieren die anfallenden Daten wiederum im faktorenanalytischen Modell (THURSTONE 1969). Die Kommunalitäten werden iterativ geschätzt, die Rotationen erfolgen teils orthogonal, teils nicht-orthogonal. Vereinzelt wird grapisch nachrotiert.

2.5.4 Ergebnisse

Zunächst zu den Befunden aus den JUNGEN-Daten:

Es zeigen sich Hinweise auf bis zu 5 bedeutsame Faktoren. Obgleich der gesamte Schiefwinkligkeitsbereich über alle in Betracht kommenden n-faktoriellen Lösungen hinweg systematisch abgesucht wird, ist unter genereller Schiefwinkligkeitsannahme der Faktoren keine theoretisch befriedigende Gesamtstruktur zu extrahieren. Dagegen führen die unter Orthogonalitätsannahme durchgeführten Berechnungen zu theoretisch gut verständlichen Strukturen.

Die varimaxrotierte Dreifaktorenlösung erbringt einen Faktor "Aggressionsmotiv" (F1), einen Faktor "Aggressions-Hemmung" (F2) sowie einen Faktor "Freundschaft" (F3). Nähere Einzelheiten sind der Tabelle 8 zu entnehmen.

Die equimaxrotierte Vierfaktorenlösung führt demgegenüber zu einer Aufspaltung des unspezifischen Aggressions-Hemmungs-Faktors (8) in einen Faktor "intrinsischer" (F3) und einen Faktor "extrinsischer Hemmung von Aggression" (F4). Die übrigen Dimensionen bleiben konstant: Aggressionsmotiv (F1) und Freundschaft (F2).

Die equimaxrotierte Fünffaktorenlösung zeigt die erwartete Unterscheidung zwischen einer Dimension generalisierter, erfolgszuversichtlich antizipierter Zielvorstellungen (F2) und einer Dimension allgemeiner Schemata von Handlungsmustern und handlungsbezogenen Kognitionen zur Verletzung (F4) innerhalb des Aggressionsmotivs einerseits und den erwarteten Split in intrinsische (F1) und extrinsische Aggressions-Hemmung (F5) andererseits. Die Freundschafts-Statements spannen in dieser Lösung den Faktor F3 auf. Einzelheiten dazu sind der Tabelle 9 zu entnehmen.

Nun zu den Befunden aus den MÄDCHEN-Daten:

Auch hier liegen Hinweise auf bis zu fünf bedeutsame Faktoren vor. Im einzelnen bestätigen die analog zu den Jungen-Daten durchgeführten Analysen die oben dargestellten Befunde hinsichtlich des Zweikomponentenmodells (varimaxrotierte Dreifaktorenlösung)

ROTIERTE FAKTOREN

VARIABLE	F1	F2	F3	h^2
01	.63	.06	-.19	.44
02	.71	-.07	-.34	.62
03	.69	.06	-.30	.57
04	.56	.22	-.11	.37
05	.76	-.03	-.25	.64
06	.79	.00	-.17	.65
07	.75	-.14	-.23	.64
08	.70	.04	-.13	.51
09	.72	.06	-.04	.52
10	.72	-.05	-.26	.59
11	.65	.01	-.09	.43
12	.56	.13	.07	.34
13	.83	-.06	-.14	.71
14	.47	.07	-.06	.23
15	-.02	.51	.36	.39
16	-.13	.61	.46	.60
17	-.17	.64	.43	.62
18	-.09	.63	.38	.61
19	-.02	.70	.32	.59
20	-.17	.64	.36	.57
21	.13	.51	.13	.29
22	-.09	.56	.50	.57
23	.28	.53	-.05	.36
24	.10	.68	.09	.48
25	-.10	.74	.22	.61
26	.29	.57	-.15	.43
27	.18	.65	.01	.46
28	.06	.75	.08	.57
29	-.15	.25	.45	.29
30	-.18	.35	.51	.42
31	-.34	.06	.71	.62
32	-.38	.13	.71	.67
33	-.18	.35	.59	.50
34	-.31	.25	.67	.61
35	-.11	.14	.64	.44
	7.41	6.10	4.45	SUMME a_j^2
	41.26	33.96	24.78	% der gemeinsamen Varianz
	41.26	75.22	100.00	kumulierte %
	21.17	17.43	12.71	% der Gesamtvarianz
	21.17	38.60	51.31	kumulierte %

Tab. 8 Ladungsmatrix der varimaxrotierten Dreifaktorenlösung (JUNGEN; n = 282; TR/SPSS, FACTOR (PA2))

ROTIERTE FAKTOREN

VARIABLE	F1	F2	F3	F4	F5	h^2
01	-.08	.37	-.18	.50	.13	.44
02	-.15	.44	-.34	.55	.02	.63
03	-.05	.50	-.30	.47	.13	.58
04	.16	.25	-.18	.52	.15	.41
05	-.12	.71	-.22	.37	.08	.71
06	-.04	.74	-.17	.39	.06	.73
07	-.10	.64	-.25	.42	-.07	.66
08	-.07	.25	-.13	.73	.09	.63
09	-.01	.38	-.06	.63	.09	.55
10	-.21	.49	-.22	.52	.10	.61
11	-.09	.31	-.08	.60	.08	.48
12	.01	.33	.08	.45	.17	.35
13	-.08	.74	-.13	.44	.01	.76
14	.13	.21	-.14	.45	.00	.29
15	.53	.02	.26	-.03	.27	.42
16	.59	-.11	.36	-.07	.34	.61
17	.65	-.16	.31	-.07	.32	.65
18	.60	-.10	.29	-.03	.42	.63
19	.60	-.12	.22	.09	.43	.62
20	.58	-.26	.26	.01	.37	.61
21	.43	.09	.06	.10	.33	.32
22	.72	-.02	.35	-.08	.20	.69
23	.29	.30	-.06	.08	.47	.40
24	.23	.03	.15	.07	.69	.56
25	.47	-.09	.21	-.07	.56	.61
26	.10	.23	-.07	.14	.65	.51
27	.17	.00	.07	.21	.69	.55
28	.35	-.01	.11	.06	.68	.60
29	.29	-.01	.42	-.16	.12	.31
30	.28	-.11	.52	-.14	.25	.44
31	.17	-.25	.72	-.19	-.03	.65
32	.23	-.20	.71	-.30	.01	.69
33	.42	-.22	.52	-.02	.14	.52
34	.30	-.29	.65	-.13	.10	.62
35	.14	-.12	.67	-.02	.09	.49
	4.02	3.90	3.85	3.98	3.58	SUMME a_j^2
	20.79	20.18	19.92	20.59	18.52	% der gemeinsamen Varian
	20.79	40.97	60.89	81.48	100.00	kumulierte %
	11.49	11.14	11.00	11.37	10.23	% der Gesamtvarianz
	11.49	22.63	33.63	45.00	55.23	kumulierte %

Tab. 9 Ladungsmatrix der equimaxrotierten Fünffaktorenlösung (JUNGEN; n = 282; TR/SPSS, FACTOR (PA2))

voll und hinsichtlich des Vierkomponentenmodells des Aggressionsmotiv-Systems weitgehend: In der equimaxrotierten Fünffaktorenlösung gelingt uns eine vergleichbare Trennung innerhalb der Aggressions-Hemmung in einen intrinsischen und einen extrinsischen Anteil, jedoch keine deckungsgleiche Differenzierung innerhalb des Aggressionsmotivs, vergleichbar dem Split aus den Jungen-Daten. Vielmehr trennen wir hier eine Dimension mehr indirekt-allgemeiner (F1) von einer Dimension mehr vordergründig-körperlicher Aggression (F5) ab. Die graphische Nachrotation des Faktors F5 um 18 Grad im und des Faktors F1 um 20 Grad entgegen dem Uhrzeigersinn zu der schiefwinkligen Faktorenkonstellation F1'/ F5' läßt diesen Differenzierungsaspekt noch deutlicher hervortreten. F1' ist nun als Dimension aggressiv-feindseliger Emotionen und F5' als Dimension körpermotorischer Aggression interpretierbar.

Die Freundschafts-Statements spannen stets einen eigenen Faktor auf, und unter genereller Schiefwinkligkeitsannahme der Faktoren durchgeführte Analysen erbringen auch bei den Mädchen ein insgesamt negatives Resultat.

2.5.5 Diskussion

Damit werden unsere im Anschluß an die Ergebnisse der Voruntersuchung formulierten Erwartungen zur Grob- und Feinstruktur der Motivierungsdaten auf Variationsebene von Statements (Hypothesen a, b) gestützt. Beide Hauptkomponentenmodelle des Aggressionsmotiv-

Systems (Zwei- und Vierkomponentenmodell) sind somit als invariant hinsichtlich unterschiedlicher Zusammensetzung der Stichprobe zur Darstellung der Referenzpopulation des AM-GITTERs (9 bis 11jährige Jungen) qualifizierbar.

Desweiteren wird unsere Erwartung bestätigt, daß - auf den gleichen Situationsbereich bezogen - freundschaftsthematische Kognitionen einen eigenen Faktor aufspannen. Die Nachweisbarkeit dieses Faktors, zumal über verschiedenfaktorielle Lösungen hinweg, bestätigt nachdrücklich die aus der Analyse von TAT-Geschichten gewonnenen Einsichten zur Ambiguität und thematischen Bandbreite des Bildersatzes und erhöht zudem, gerade auch deshalb, die interne Validität der oben referierten Befunde.

Gehen wir etwas mehr ins Detail!

Es zeigt sich also, daß die internen Zusammenhänge zwischen Testdaten des AM-GITTERs , pro Motivindikator über den situationsseitigen Bestandteil des Verfahrens hinweg ermittelt, auch im faktorenanalytischen Modell durch die Konstrukta "Aggressionsmotiv" und "Aggressions-Hemmung" beschreibbar und aus ihnen heraus erklärbar sind. Das wird im Wiederholungsversuch an einer zweiten, unabhängigen Stichprobe bestätigt. Die Ergebnisse, die wir auf der Grundlage einer Mädchen-Stichprobe errechnet haben, stützen diesen Befund zusätzlich. Letzteres ist deshalb ermutigend, weil das Verfahren in spezifischen Aspekten seiner Gestaltung für diese Bezugsgruppe garnicht vorgesehen ist und diese daher unter "härteren" Bedingungen gemessen wurde.

Motivationstheoretisch handelt es sich bei den Faktoren des Modells zweier Hauptkomponenten um fundamentale Dimensionen. Empirisch wird das dadurch bestätigt, daß der in ihnen enthaltene Gliederungsgesichtspunkt unabhängig vom Differenziertheitsgrad einer faktoriellen Struktur relevant bleibt. Es lassen sich stets zwei Faktorgruppen unterscheiden, von denen die eine Aggressionsmotiv-, die andere Aggressions-Hemmungs-Faktoren umfaßt. Damit können wir an Ergebnisse anknüpfen, die auf ähnliche Weise in anderen Handlungsbereichen, auf die eine (allgemeine) Theorie der Motivation angewandt wurde, zutage traten. So haben sich motivationale Dimensionen vom Typ der Hoffnung und der Furcht in faktoriellen Analysen von Meßdaten zum Leistungsmotiv-System (SCHMALT 1976 a, S. 106 ff.; HECKHAUSEN 1980, S. 257) und neuerdings auch in solchen zum Machtmotiv-System (SCHMALT 1979) extrahieren lassen. Spezielle Aufmerksamkeit verdient in diesem Zusammenhang das 6-faktorielle Klassifikationsschema für Motivkomponenten, das SCHMALT (1976 a, S. 118) für das Leistungsmotiv-System auf der Grundlage einer fünffaktoriellen Rotationslösung vorgestellt hat und das als sechsten Faktor eine hypothetische Komponente enthält. Auch dort werden die Faktoren durch die o.g. Gesichtspunkte zentral untergliedert.

Bemerkenswert ist ferner, daß sich der motivationstheoretisch (vorläufig allerdings mehr im Sinne einer Arbeitshypothese) postulierte Orthogonalitätscharakter unter diesem Gesichtspunkt errechneter Zweifaktorensysteme der Aggressivität gut hat bestätigen lassen. Wir hatten zwar mit unseren selektiven Maßnahmen bei

der Zusammenstellung der zweiten Testvorform die Wahrscheinlichkeit, nochmals orthogonale Strukturen zu erhalten, maximiert. Dem steht jedoch entgegen, daß mit dem gleichen Verfahren in einer anderen Population, den Mädchen, auf Anhieb orthogonale Strukturen ermittelt wurden. Allein schon aufgrund andersgearteter Sozialisierungsbedingungen, unter denen sich deren Motivsystemgenese vollzieht (vgl. FESHBACH 1970, S. 188 ff.; BARON 1977, S. 218 ff.; ZILLMANN 1979, S. 347 ff.; MACCOBY 1980, S. 216 f.), hätte man dort mit abweichenden Ergebnissen rechnen können.

Es wird jedoch auch der Nachweis geführt, daß man die gleichen internen Zusammenhänge zwischen den Testdaten zu einem noch höheren Prozentanteil aufklären und beschreiben kann, wenn man in einem höherfaktoriellen Modell auf differenziertere motivationstheoretische Konstrukta rekurriert. Man gerät dadurch keineswegs in einen Widerspruch zu der oben diskutierten Befundlage, da ja einerseits, wie schon gesagt, der in der zweifaktoriellen Hauptkomponentenstruktur enthaltene zentrale Differenzierungsgesichtspunkt erhalten bleibt und andererseits diese differenzierteren Konstrukta gerade theoretisch postulierten weiteren Teilkomponenten des Aggressionsmotiv-Systems entsprechen (vgl. KORNADT 1982 a, S. 273 ff.; 1982 c, S. 95 ff.). Dabei ist natürlich zu beachten, daß zunehmend spezifischere Dimensionen das Aggressionsmotiv-System in immer enger umgrenzte Teilbereiche ordnen und wegen der damit verbundenen progressiven Abnahme ihres Beziehungsreichtums an allgemeinem Erklärungs- und Stellenwert verlieren.

Daß die Freundschafts-Statements in allen diskutierten Lösungen einen eigenen Faktor aufspannen, trägt zunächst, wie schon erwähnt, zur internen Validität der Befunde bei. Da es sich bei dieser Dimension jedoch darüber hinaus um ein für die Situationsklasse des AM-GITTERs wichtiges extrinsisches Gegenmotiv zum Aggressionsmotiv handelt, läßt ihre Existenz die Annahme an Wahrscheinlichkeit gewinnen, daß wir mit unserem Faktor "extrinsische Aggressions-Hemmung" den aggressionsthematisch-fixierten Teil der im übrigen relativ offenen Klasse der gesamten extrinsischen Hemmung umschrieben haben, wie das von der Konstruktion her ja auch vorgesehen war.

Wenn man die graphisch nachrotierte Fünffaktorenlösung, die wir auf der Grundlage der Mädchen-Daten errechnet haben, mit der korrespondierenden Fünffaktorenlösung in der Jungen-Stichprobe vergleicht, stellt man fest, daß der Aggressionsmotiv-Split in beiden Gruppen von Kindern gleichermaßen zu einem Faktor handlungsbezogener Kognitionen einerseits sowie zu einem zweiten Faktor andererseits führt, der einem weiteren Teilbereich des Motivs zuzuordnen ist. Während sich dieser bei den Jungen mehr in kognitiv formulierten Zielvorstellungen verselbständigt, tritt bei den Mädchen das affektive Moment faktoriell in den Vordergrund. Dabei ist zu beachten, daß dieser Unterschied in dem Teil des Motivsystems auftritt, der - nach unseren bislang vorliegenden Erfahrungswerten - quasi "bildabhängig" ist, denn die LESSER-Bilder besitzen eine primäre Anregungsqualität für das Aggressionsmotiv und für die Aggressions-Hemmung lediglich eine sekundäre. Der situationsseitige Bestandteil

des AM-GITTERs ist aber nun gerade der, der u.E. bei einer Adaptation des Verfahrens für Mädchen in weit beträchtlicherem Umfang geändert werden müßte als der personseitige, sodaß man bei einer in beiden Populationen äquivalenten Messung erwarten könnte, daß dieser Unterschied ggfls. wieder in den Hintergrund tritt; jedenfalls ist das nicht auszuschließen.

Die Befunde bestätigen somit die Schlußfolgerungen, die wir aus den vergleichbaren Analysen der Voruntersuchung gezogen haben, noch deutlicher als erwartet. Man kann sagen, daß sie für eine gelungene motivationstheoretische Ableitung des Verfahrens sprechen und darüber hinaus konkrete Ansatzpunkte zur Formulierung von Meßzielen bieten. Da wir von einem gleichberechtigten Nebeneinander beider Modellalternativen (zwei- und vierfaktorielles Hauptkomponentenmodell) ausgehen können, werden wir auch auf Grundlage beider Skalen erstellen und im übrigen den Versuch machen, die Binnen- als auch die Querbeziehungen zwischen diesen strukturell und inhaltlich aufzuklären.

Zuvor jedoch wollen wir die Daten auf die Möglichkeit zur Etablierung eines Generalfaktormodells prüfen. Diesem Versuch soll der nächste Abschnitt gewidmet werden.

2.5.6 Generalfaktormodell

Wie schon in der Voruntersuchung gehen wir wieder von der varimaxrotierten Zweifaktorenlösung aus. Der Faktor F1 ist hier als "Aggressionsmotiv", der Faktor F2 als "Aggressions-Hemmung" interpretierbar. Von Interesse sind nun die Ladungsvektoren der F-Statements in diesem Modell. Sie laden F1 negativ und gleichzeitig F2 positiv. Hier wird nun etwas deutlich, was in der dreifaktoriellen Rotationslösung nicht klar erkennbar wurde, nämlich, daß sie offensichtlich mit beiden Dimensionen etwas gemeinsam haben, jedoch in einmal gleicher, dann umgekehrter Richtung. Es ist klar, daß die rotierte Zweifaktorenlösung in diesem Zustand noch nicht endgültig befriedigt. Wir versuchen nun ohne die Extraktion eines dritten Faktors und anschließende Rotation, was, wie wir ja bereits wissen, die oben angedeutete Gemeinsamkeit wieder völlig in den Hintergrund treten lassen würde, weiterzukommen, indem wir die varimaxrotierte Zweifaktorenlösung in einen besser interpretierbaren Zustand überführen.

Die graphische Darstellung der o.g. Zweifaktorenlösung liefert dazu wieder einen markanten Hinweis, ein durch die F-Statements rechtwinklig abgewinkeltes "schmales Band", das die Endpunkte der Ladungsvektoren bilden. Wir nutzen diesen Hinweis zu einer graphischen Nachrotation des Systems um 47 Grad im Uhrzeigersinn. Danach befindet sich die Lösung in einem Zustand, der das bereits vorgestellte Generalfaktormodell eindeutig bestätigt. Den Faktor

F2' können wir als thematischen Bereichsfaktor AGGRESSION, den Faktor F1' als Dimension der SOZIALEN WERTUNG identifizieren. Die F-Statements laden die beiden Dimensionen wie theoretisch angenommen.

Wie schon in der Voruntersuchung, so liegt auch hier das Generalfaktormodell bereits der unrotierten Zweifaktorenlösung zugrunde. In der Voruntersuchung hatten wir eine Nachrotation um 41 Grad UZS vorgeschlagen, hier erwies sich eine Nachrotation um etwa 47 Grad UZS als optimal. Wir entnehmen daraus, daß durch die VARIMAX-Rotation eine Drehung des unrotierten Zweifaktorensystems um etwa 45 Grad entgegen dem Uhrzeigersinn bewirkt wird. Damit ist der Übergang vom (unrotierten) Generalfaktormodell des Aggressionsmotiv-Systems zum Zweikomponentenmodell (sensu KORNADT) empirisch beschrieben. Theoretisch kann man das (umgekehrt) leicht nachvollziehen. Man definiert die erste Winkelhalbierende im Zweikomponentenmodell als thematischen Faktor AGGRESSION und gliedert daran die bipolare Dimension SOZIALE WERTUNG orthogonal an. Im neuen System sind die erstmals beschreibenden Größen nunmehr beschriebene. Es bietet sich deshalb an, im folgenden, auch aus Gründen der Objektivität, beim Generalfaktormodell nach Möglichkeit von der unrotierten Zweifaktorenlösung auszugehen.

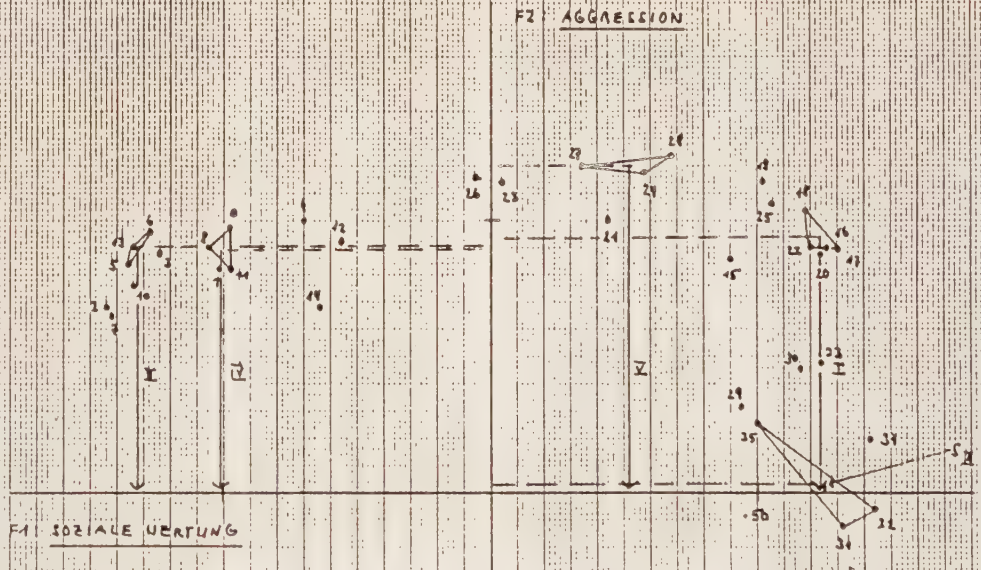
Wir suchen nun nach Hinweisen auf integrative Eigenschaften des Generalfaktormodells und bedienen uns dabei der bereits in der Voruntersuchung beschriebenen Methode (vgl. Abschnitt 1.3.5.5 Teil II). Speziell versuchen wir, die Faktoren der equimaxrotierten

Fünffaktorenlösung (vgl. Tab. 9) im Generalfaktormodell zu markieren. Die Einzelheiten dazu sind Abbildung 6 zu entnehmen.

Wir prüfen nun die Integrations- und Leistungsfähigkeit des Generalfaktormodells noch einmal auf Grundlage der von den Mädchen erhaltenen Motivierungsdaten. Dazu verwenden wir die in Abschnitt 2.5.4 vorgestellte equimaxrotierte Fünffaktorenlösung (ohne graphische Nachrotation). Die Ergebnisse sind Abbildung 7 zu entnehmen.

Damit ist nun auch unsere Ausgangshypothese (c), die Integrierbarkeit unterschiedlicher Varianten von Hauptkomponentenmodellen im unrotierten zweifaktoriellen Generalfaktormodell, als bestätigt anzusehen, und zwar im Wiederholungsversuch auf Grundlage einer neu gezogenen Jungen-Stichprobe und darüber hinaus erstmalig auch auf Grundlage von Mädchen erhaltener Motivierungsdaten.

Wir wollen nun noch auf ein Ergebnis zu sprechen kommen, das uns in dem o.g. Zusammenhang nicht ganz unwesentlich zu sein scheint. Wie wir bereits wissen, erzeugen wir bei Extraktion von drei Faktoren und anschließender VARIMAX-Rotation ein Dreikomponentenmodell, bestehend aus den Faktoren "Aggressionsmotiv", "Aggressions-Hemmung" und "Freundschaft". Unterlassen wir allerdings die Rotation, so läßt die unrotierte Dreifaktorenlösung die Vorstellung zu, an das vorgestellte 2-dimensionale Generalfaktormodell einen zweiten unipolaren Faktor orthogonal anzugliedern, den wir als thematischen Bereichsfaktor "FREUNDSCHAFT" interpretieren können. Die Ladungen auf dieser Dimension sind jedoch nicht sehr ausgeprägt, was jedoch bei den



Nr. 1 – 14 : A-STATEMENTS
15 – 28 : A# ——— ———
29 – 35 : # ——— ———

I – V: MARKIERVARIABLENCLUSTER DER EQUIMAXROTIERTEN 5-FL (VGL. FAB 9)

I ENTRINSISCHE AGGRESSIONS-HEMMUNG
II BEDÜRFNIS UND ERFOLGSERWARTUNG AGGRESSION
III FREUNDSCHAFT
IV AGGRESSIONS-HANDLUNG
V EXTRINSISCHE AGGRESSIONS-HEMMUNG

Abb. 6 Ladungskoordinaten der Statements des AM-GITTERS in der unrotierten Zweifaktorenlösung und integrative Aspekte des Generalfaktormodells (Hauptuntersuchung, GAUSEBU, n = 282; IFK/SPSS; FACTOR (PA2))

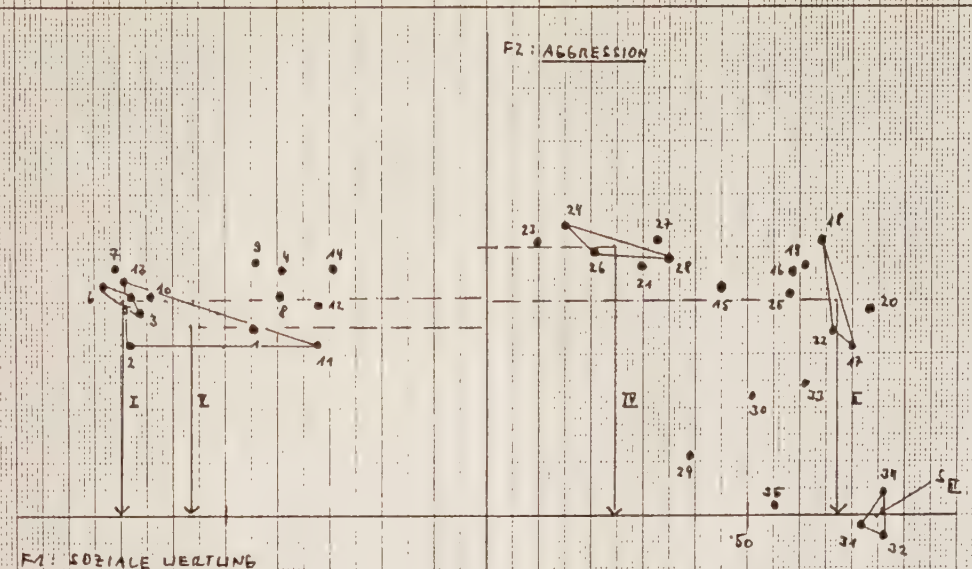


Abb. 7: Ladungskoordinaten der Statements des AM-GITTERS in der unrotierten Zweifaktorenlösung und integrative Aspekte des Generalfaktormodells (Hauptuntersuchung; MÄDCHEN, n = 284; TR/SPSS, FACTOR (PA2))

Bedingungen, unter denen wir die F-Statements entwickelt und faktorisiert haben, nicht verwundert. Für die Erweiterung der motivthematischen Bandbreite des AM-GITTERs um ein nicht aggressionsthematisches, soziogenes Gegenmotiv (FREUNDSCHAFT), könnte dieser Faktor jedoch modellrelevant werden. Um die gemeinsame "Achse" SOZIALE WERTUNG, die nun Generalfaktorstatus besitzt, könnte man die spezifischen thematischen Bereichsfaktoren AGGRESSION und FREUNDSCHAFT herum anordnen. Die Hoffnungs- und Furcht-Komponenten des Aggressionsmotiv-Systems würden sich dann in der XY-Ebene, die des Freundschaftsmotiv-Systems in der XZ-Ebene beschreiben lassen (vgl. Abb. 8).

Dieser letztgenannte Befund wird auch durch die an den Mädchen erhobenen Daten tendenziell gestützt.

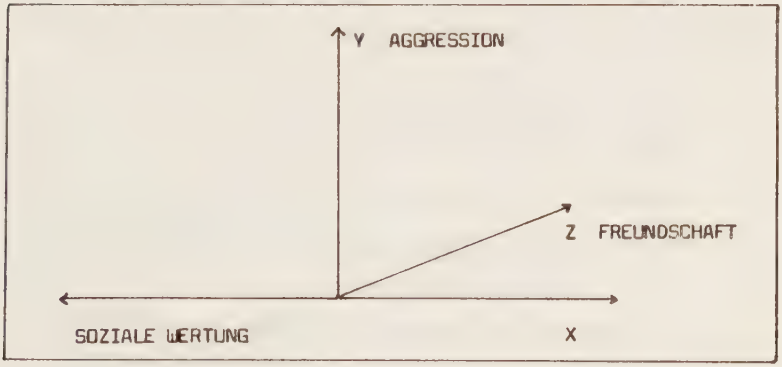


Abb. 8 Dreidimensionales Generalfaktormodell zur integrierten Darstellung zweier soziogener Motive, AGGRESSION und FREUNDSCHAFT (hypothetisch) auf Grundlage der unrotierten Dreifaktorenlösung (JUNGEN, n = 282; TR/FACTOR (PA2))

2.6 Selektion der endgültigen Skalen

2.6.1 Vorüberlegungen und Wahl des Modells

Durch die Dimensionierungsanalysen der Statements konnten wir die drei Modellvarianten, die sich uns in der Voruntersuchung zur Entwicklung von Skalen anboten, alle gleichermaßen gut bestätigen. Die eleganteste Lösung wäre es nun zweifelsohne, die Skalen direkt aus dem Generalfaktormodell abzuleiten. Wir wissen jedoch noch nicht, ob sich dieses auch im Wiederholungsexperiment auf Subskalen-Ebene bestätigen läßt. Aus diesem Grund halten wir es für vernünftig, die Skalen direkt aus den beiden Hauptkomponentenmodellen abzuleiten und deren Dimensionalität anschließend erneut zu prüfen.

Die nachfolgend berechneten Skalen teilen wir demzufolge in 2 Gruppen ein. Maßgeblich für diese Unterteilung sind also die beiden Modelle, die ihrer Konstruktion zugrunde liegen.

Die erste Gruppe von Skalen wollen wir "Standard-Skalen" (STD-SKALEN) nennen, da sie der Standard-Auswertung des A-TAT entsprechen. Es handelt sich dabei zunächst um die beiden Skalen, mit denen wir das Aggressionsmotiv und die Aggressions-Hemmung direkt (fundamental) messen, und im weiteren um die davon abgeleiteten Skalen für Netto-Aggression und Gesamt-Aggressivität. Modellvorlage für die Entwicklung dieses Sets ist das zweifaktorielle Hauptkomponentenmodell des Aggressionsmotiv-Systems, wie es der varimaxrotierten Dreifaktorenlösung (vgl. Tab. 8) zugrundeliegt.

Die zweite Gruppe von Skalen wollen wir "Differentielle Skalen" (DIFF-SKALEN) nennen und mit ihnen differentielle Aspekte des Aggressionsmotivs und der Aggressions-Hemmung erfassen, die wir faktorenanalytisch belegt haben. Modellvorlage des zweiten Sets von Skalen ist das vierfaktorielle Hauptkomponentenmodell des Aggressionsmotiv-Systems, wie es der equimaxrotierten Fünffaktorenlösung (vgl. Tab. 9) entspricht. Daraus resultieren also vier weitere Skalen.

Im Generalfaktormodell sollte sich die "Eigenständigkeit" der 8 selektierten Skalen an einer unterscheidbaren Zeigerstellung ihrer Ladungsvektoren ablesen lassen, falls wir es auf Subskalen-Ebene bestätigen können.

2.6.2 Richtlinien der Zusammenstellung

Die fundamentalen Skalen beider Modelle wollen wir unter folgenden Zielvorgaben und Richtlinien für die Statements zusammenstellen:

- gleiche Itemzahl für die Skalen innerhalb eines Modells;
- bestmögliche Rohwerteverteilung;
- maximale Reliabilität;
- minimale Interdependenz;
- maximale Ladung der vorgesehenen Dimension;
- minimale Ladung nicht vorgesehener Dimensionen;
- faktorielle Komplexität $K < 2$;
- faktorielle Stabilität.

Zu einzelnen dieser Punkte möchten wir jedoch, bevor wir fortfahren, noch einige Anmerkungen machen. Ad Reliabilität: wir berücksichtigen nicht nur Schätzungen der internen Konsistenz,

sondern auch die mit jeder Selektion verbundenen Schätzungen der Retest-Reliabilitäten, über die wir jedoch erst später berichten (vgl. Abschnitt 3.1). Wir ermitteln die Stabilitätsschätzungen stets für Retest-Intervalle von 2, 4 und 8 Wochen. Ad minimale Ladung nicht vorgesehener Dimensionen: Wir unterscheiden bei den DIFF-Skalen zwischen konstruktnahen und konstruktfernen Dimensionen und entscheiden im Zweifelsfall zugunsten einer Ladung konstruktnaher Dimension. Ad faktorielle Komplexität: Wir bestimmen die faktorielle Komplexität K eines Statements als Anzahl der Ladungen, für die gilt: $a^2 : h^2 \geq .50$. Für die fundamentalen STD-Skalen sollte nach Möglichkeit ein Wert von K = 1 gegeben sein (und zwar bei ausreichender Kommunalität); da wir aus diesem Vorrat die Statements für die DIFF-Skalen auswählen, akzeptieren wir für letztere auch einen K-Wert von 0 (im zugehörigen Modell). Ad faktorielle Stabilität: Wir beziehen die faktorielle Stabilität einer Statement-Variablen als Güte-Kriterium in die Zusammenstellung der Skalen mit ein, um faktorielle Invarianz der Meßwerte zu gewährleisten. Zum Aufspüren "faktorieller Sprünge" einer Statement-Variablen dient der Vergleich ihrer Ladungsvektoren aus Vor- und Hauptuntersuchung. Aus einer Vielzahl von Gründen halten wir dabei die aus der Hauptuntersuchung bekannten Koordinaten der Ladungsvektoren für aussagekräftiger, zumal letztere ja unter wesentlich günstigeren experimentellen Bedingungen durchgeführt wurde, worauf auch die Ergebnisse insgesamt verweisen (9).

Im Gegensatz zur Voruntersuchung gehen wir nun nicht mehr davon

aus, daß zwei DIFF-Skalen sich additiv zu einer STD-Skala zusammenfassen lassen, denn unter diesem Aspekt wären in beiden Modellen nicht mehr die optimalen Skalen berechenbar. Zudem verfügen wir ja jetzt schon über eine Modellvorstellung, die die Zusammenhänge zwischen STD- und DIFF-Skalen beschreibt.

Für die einzelnen Selektionen wählen wir jeweils alle 6 in die Hauptuntersuchung einbezogenen Bildsituationen. Für die STD-Skalen können wir sie aufgrund der in Abschnitt 2.4.3 dargestellten Ergebnisse beibehalten. Für die DIFF-Skalen liegen jedoch zunächst noch keine vergleichbaren Ergebnisse vor. Wir schätzen daraufhin die Bildsituationen nach der gleichen Methode für die einzelnen Konstrukta des Vierkomponentenmodells erneut ein. Die Ergebnisse sprechen für die Beibehaltung aller 6 Bildsituationen.

Nun rechnen wir unter den o.g. Vorgaben eine Reihe von Probeselektionen durch, bei denen wir einzelne Aspekte jeweils etwas unterschiedlich gewichten. Im Rahmen dieser Selektionen bestimmen wir jeweils auch sämtliche Parameter für die abgeleiteten STD-Skalen. Nach diesen insgesamt recht umfangreichen Analysen entscheiden wir uns für die nachfolgend vorgestellte Zusammenstellung des AM-GITTERs und seiner (fundamentalen) Skalen.

2.6.3 Standard-Skalen

2.6.3.1 Fundamentale Skalen

Zur Erstellung einer Aggressionsmotiv-Skala A scheiden wir aus der zu diesem Zweck geprüften Gruppe von Statements (01 bis 14; vgl. auch im folgenden stets ANHANG II) die Statement-Variable Nr. 14 aus und behalten die übrigen (01 bis 13) bei.

Zur Erstellung einer Aggressions-Hemmungs-Skala AH scheiden wir aus der dafür entwickelten Statement-Gruppe (15 bis 28) die Statement-Variable Nr. 21 aus, alle übrigen (15 bis 20; 22 bis 28) behalten wir bei.

Die Rohwerteverteilungen der beiden fundamentalen STD-Skalen sind in Abbildung 9 dargestellt. Weitere Verteilungscharakteristika haben wir im ANHANG IV zusammengestellt (gilt jeweils auch für die nachfolgend vorgestellten Skalen).

2.6.3.2 Abgeleitete Skalen

Entsprechend dem Vorgehen von KORNADT (1982 b, S. 28) leiten wir aus den fundamentalen STD-Skalen zwei weitere STD-Skalen durch Summen- und Differenzenbildung ab. Und zwar berechnen wir die Skala für das Konstrukt Netto-Aggression, n-A., nach der Formel n-A. = A - AH und die Skala für das Konstrukt Gesamt-Aggressivität, Ges.-A., nach der Formel Ges.-A. = A + AH. Zu den Rohwerteverteilungen der abgeleiteten STD-Skalen vgl. Abbildung 10 (vgl. auch ANHANG IV).

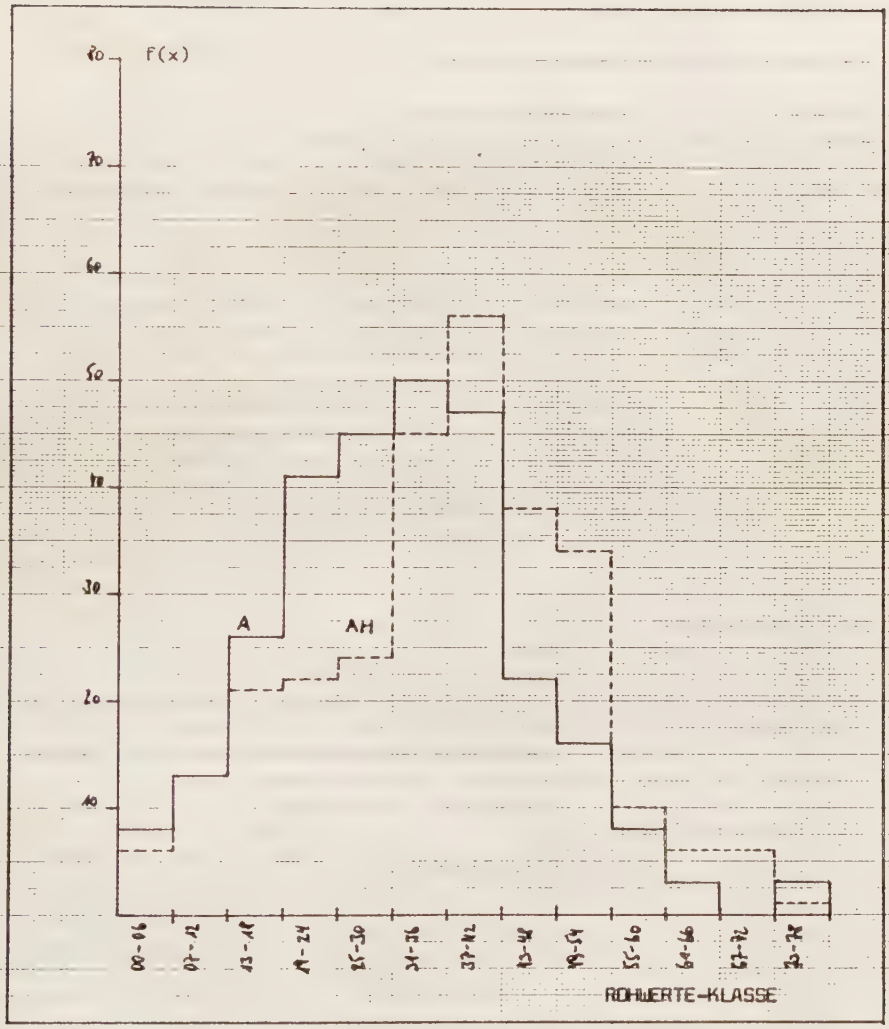


Abb. 9 Verteilung der A- und der AH-Rohwerte des AM-GITTERs.
(Hauptuntersuchung, JUNGEN, n = 282; SI/SPSS, FREQUENCIES)

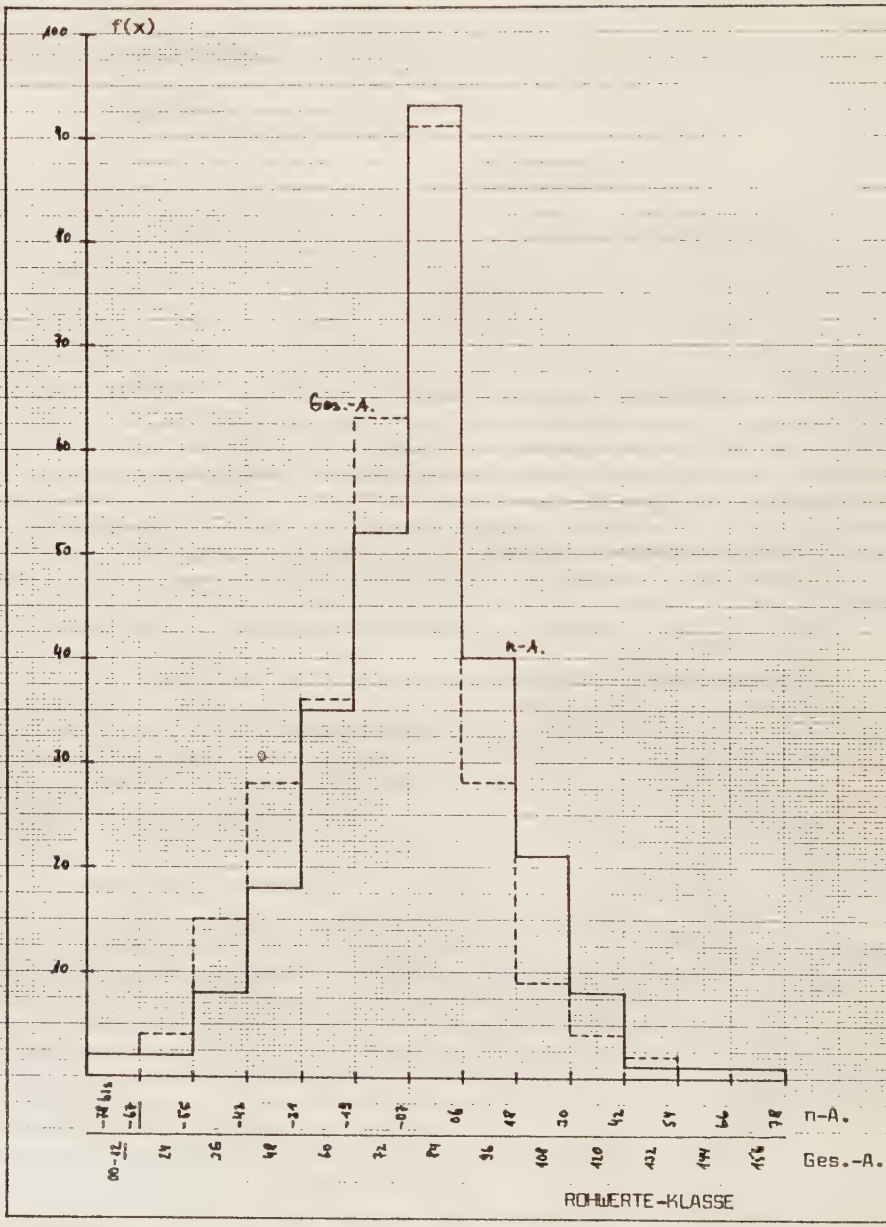


Abb. 10 Verteilung der n-A.- und Ges.-A.-Rohwerte des AM-GITTERs.
(Hauptuntersuchung, JUNGEN, n = 282; SI/SPSS, FREQUENCIES)

2.6.4 Differentielle Skalen

Mit der DIFF-Skala ADM versuchen wir den Aspekt des Aggressionsmotivs meßbar zu machen, der sich insbesondere auf das Bedürfnis zu verletzen und die Hoffnung auf Aggressionserfolg bezieht. Wir wählen dazu aus unserem Statement-Vorrat die Nummern 03, 05, 06, 07 und 13 aus.

Mit der DIFF-Skala ADH versuchen wir einen weiteren Aspekt des Aggressionsmotivs, handlungsbezogene Kognitionen zur Verletzung, zu erfassen. Wir wählen dazu die Statements mit den Nummern 08, 09, 10, 11 und 12.

Die Rohwerteverteilungen der DIFF-Skalen ADM und ADH sind in der Abbildung 11 dargestellt.

Mit der DIFF-Skala AHI versuchen wir die intrinsischen Anteile der Aggressions-Hemmung (Furcht vor negativen inneren Folgen von Aggression) zu messen. Wir wählen dazu die Statements mit den Nummern 16, 17, 18, 19 und 22.

Mit der DIFF-Skala AHE wollen wir demgegenüber die extrinsischen Anteile der Aggressions-Hemmung (Furcht vor negativen äußeren Folgen von Aggression) erfassen. Wir wählen dazu die Statements mit den Nummern 24, 25, 26, 27 und 28 aus.

Die Rohwerteverteilungen der DIFF-Skalen AHI und AHE sind in der Abbildung 12 dargestellt.

Weitere Einzelheiten zu den Meßwerteverteilungen der DIFF-Skalen vgl. ANHANG IV.

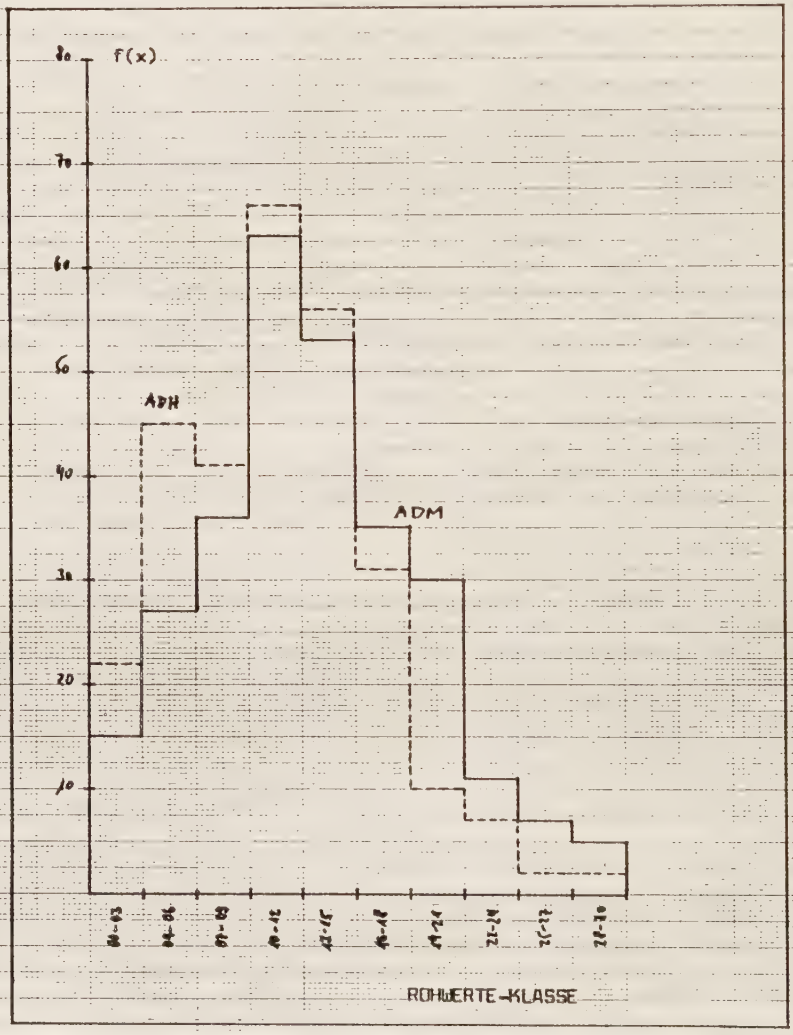


Abb. 11 Verteilung der ADM- und ADH-Rohwerte des AM-GITTERs.
(Hauptuntersuchung, JUNGEN, n = 282; SI/SPSS, FREQUENCIES)

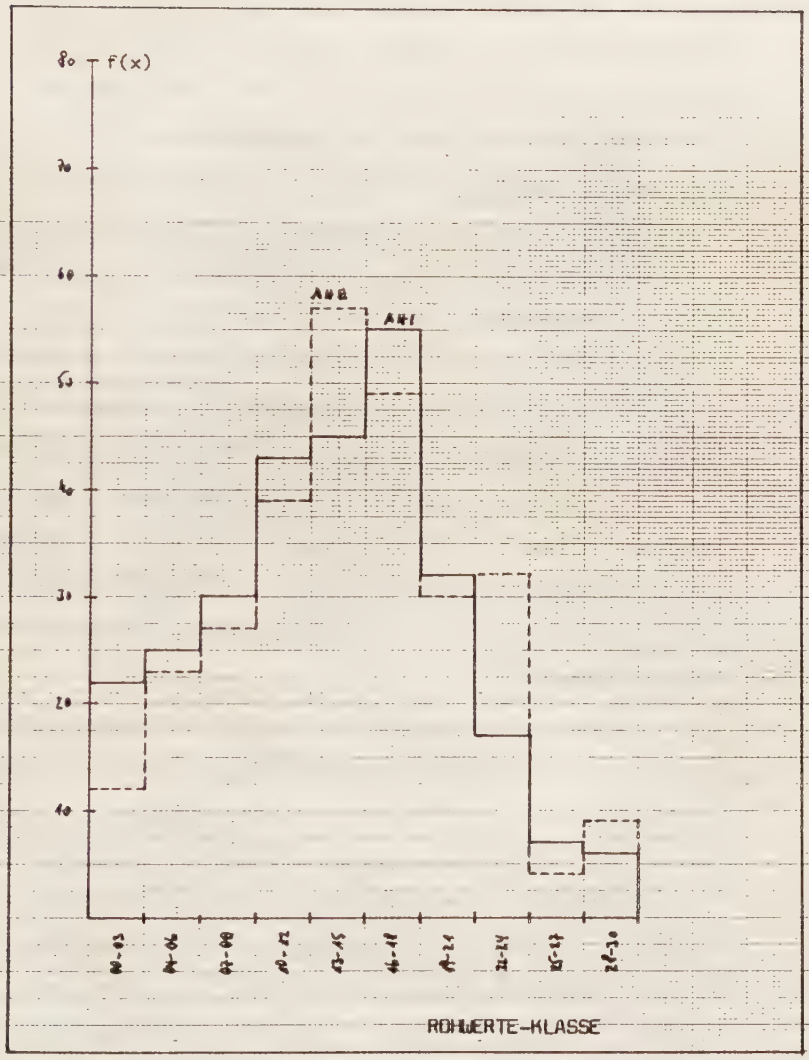


Abb. 12 Verteilung der AHI- und AHE-Rohwerte des AM-GITTERs.
(Hauptuntersuchung, JUNGEN, n = 282; SI/SPSS, FREQUENCIES)

2.7 Reliabilität der Skalen des AM-GITTERs

Wir bestimmen nun die Reliabilität der fundamentalen Skalen des AM-GITTERs nach der Methode "split-half" und werten sie anschließend nach der Formel von SPEARMAN-BROWN auf. Als zweite Methode wählen wir die Formel KR 20 von KUDER & RICHARDSON. CRONBACH (1970) hat zur Bestimmung der Reliabilität als interner Konsistenz einen ALPHA-Koeffizienten vorgeschlagen, der als Mittelwert aller möglichen split-half-Koeffizienten interpretiert werden kann. Dieser Koeffizient entkräftet die gelegentlich geäußerte Kritik der Abhängigkeit der Höhe des split-half-Koeffizienten von der speziellen Art des "splitting". Durch den Umstand bedingt, daß es sich bei unseren Items um 0/1-gescorte Meßelemente handelt, trifft für diese die Bedingung zu, daß mit KR 20 ein Spezialfall des ALPHA-Koeffizienten vorliegt. Somit sind die berechneten KR 20 - Werte auch als ALPHA-Koeffizienten interpretierbar.

Für die abgeleiteten Skalen (n-A., Ges.-A.) kommt eine Reliabilitätsbestimmung als Konsistenzschätzung nicht infrage. Dagegen läßt sich deren Reliabilität zum Zeitpunkt der Messung unproblematisch als Verhältnis der Varianz der wahren zur Varianz der beobachteten Werte definieren und somit nach den Regeln zur Berechnung der Reliabilität linear abgeleiteter Testwerte (vgl. TACK 1973, S. 21 ff.) ermitteln. Wir legen diesen Berechnungen die KR 20 - Werte der Skalen A und AH zugrunde und bestimmen deren Interkorrelation als Rangkorrelation.

Die Ergebnisse der Reliabilitätsbestimmung sind der Tabelle 10 zu entnehmen.

SKALA	split-half SB	KR 20
A	.96	.92
AH	.94	.92
ADM	.91	.85
ADH	.84	.81
AHI	.90	.86
AHE	.88	.85

SKALA	def. Varianzenverhältnis
n-A.	.95
Ges.-A.	.95

Tab. 10 Reliabilität der Skalen des AM-GITTERs, bestimmt als interne Konsistenz und als definiertes Varianzenverhältnis. (Hauptuntersuchung, JUNGEN, n = 282; TR/IT01)

Die in Tab. 10 dargestellten Ergebnisse lassen den Schluß zu, daß die STD-Skalen über eine gute, die DIFF-Skalen über eine befriedigende Meßgenauigkeit verfügen. Die etwas geringere Meßgenauigkeit der DIFF-Skalen ist auf den Umstand zurückzuführen, daß ihre Meßwerte jeweils nur auf Grundlage von 30 Items ermittelt werden; dagegen umfassen die fundamentalen STD-Skalen jeweils 78 Items und bieten damit eine wesentlich bessere Voraussetzung für hohe Meßgenauigkeit.

2.8 Dimensionalität der Skalen des AM-GITTERs

2.8.1 Ausgangshypothesen

Wir untersuchen nun die Meßwerte der 8 Subskalen der Testendform des AM-GITTERs im faktorenanalytischen Modell. Wir erwarten,

(a) daß sich die Meßwerte innerhalb des zweifaktoriellen, unrotierten Generalfaktormodells dimensional vollständig aufklären lassen;

(b) daß die Testvektoren fundamentaler Skalen für das Aggressionsmotiv ($\overrightarrow{A}$, $\overrightarrow{ADM}$, $\overrightarrow{ADH}$) einerseits und für die Aggressions-Hemmung ($\overrightarrow{AH}$, $\overrightarrow{AHI}$, $\overrightarrow{AHE}$) andererseits dabei in einander gegenüberliegende Quadranten hineinstehen, die sich durch Unterteilung des Feldes durch den unipolaren Faktor AGGRESSION ergeben, d.h. also den Faktor SOZIALE WERTUNG entgegengesetzt laden;

(c) daß die unter (b) genannten Subtests den unipolaren Faktor AGGRESSION alle substantiell laden;

(d) daß sich das Subtestvektorenbündel in folgender Anordnung in der Ebene entfaltet:
$\overrightarrow{n\text{-}A.}$, $\overrightarrow{ADM}$, $\overrightarrow{A}$, $\overrightarrow{ADH}$, $\overrightarrow{Ges.\text{-}A.}$, $\overrightarrow{AHE}$, $\overrightarrow{AH}$, $\overrightarrow{AHI}$.

Sollten sich die Hypothesen (a) bis (d) im einzelnen bestätigen lassen, so weist das Generalfaktormodell auch auf Subskalen-Ebene die integrativen Eigenschaften auf, die wir bereits auf Statement-Ebene nachgewiesen haben (vgl. Abb. 6). Aufgrund günstigerer Bedingungen bei der Durchführung der Messungen erwarten wir zudem, daß sich die o.g. Hypothesen noch deutlicher bestätigen lassen, als im vergleichbaren Fall der Voruntersuchung.

2.8.2 Methode

Wir berechnen die Matrix der Rangkorrelationen zwischen den Rohwerten der 8 Skalen und untersuchen diese dann im faktorenanalytischen Modell (THURSTONE 1969). Die Kommunalitäten werden iterativ geschätzt, und wir versuchen, ohne Nachrotation auszukommen. Der Übersicht halber stellen wir anschließend die Ladungsmatrix noch einmal graphisch dar.

2.8.3 Ergebnisse

Die Interkorrelationen der Subskalen des AM-GITTERs sind der Tabelle 11 zu entnehmen.

rho	A	ADM	ADH	AH	AHI	AHE	n-A.	Ges.-A.
A		.93	.90	-.04	-.18	.13	.64	.66
ADM			.74	-.11	-.24	.05	.66	.56
ADH				.03	-.10	.16	.53	.65
AH					.89	.86	-.74	.66
AHI						.61	-.79	.47
AHE							-.54	.70
n-A.								-.06
Ges.-A.								

Tab. 11 Korrelationsmatrix der Subskalen der Testendform des AM-GITTERs. (Hauptuntersuchung, JUNGEN, n = 282; TR/SPSS, NONPAR CORR)

Der Scree-Test verweist deutlich auf 2 extraktionswürdige Faktoren, wie Abbildung 13 zeigt:

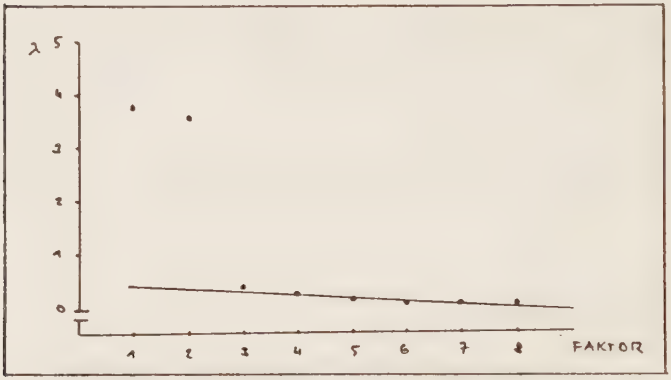


Abb. 13 Eigenwerte der rho-Matrix, der Größe nach geordnet und SCREE-TEST (n=282)

Wir extrahieren daraufhin die beiden ersten Faktoren und erhalten nach genau 1 Iteration die in Tabelle 12 dargestellte unrotierte Ladungsmatrix.

Nun ordnen wir die einzelnen Subtests auf Grundlage der in Tab. 12 angegebenen Koordinaten graphisch in die zugehörige Fläche ein und erzeugen so eine vektorgeometrische Repräsentation der Ladungsmatrix (vgl. Abb. 14).

Aus Abbildung 14 ist ersichtlich, daß sich unsere eingangs aufgestellten Hypothesen (a) bis (d) durch Berechnung klar bestätigen lassen.

	UNROTIERTE FAKTOREN		
VARIABLE	Faktor 1	Faktor 2	h^2
A	.77	.63	.99
ADM	.76	.52	.86
ADH	.65	.62	.81
AH	-.66	.74	.98
AHI	-.71	.55	.82
AHE	-.44	.76	.77
n-A.	.96	-.15	.95
Ges.-A.	.08	.96	.92
Summe a_j^2	3.66	3.43	
% gemeinsame Varianz :	51.70	48.30	
% Gesamt-Varianz :	45.75	42.88	SUMME: 88.63

Tab. 12 Ladungsmatrix der Subskalen der Testendform des AM-GITTERs. (Hauptuntersuchung, JUNGEN, n = 282; SI/SPSS, FACTOR (PA2))

Die Testendform des AM-GITTERs zur Messung verschiedener Aspekte des Aggressionsmotiv-Systems bei Kindern erweist sich somit als 2-dimensional. Bei Faktor 2 handelt es sich um eine Dimension, die der thematischen Kategorie des Aggressionsmotiv-Systems entspricht. Meßwerte zur Erfassung diverser Aspekte feindseliger Aggressivität grenzen sich durch Korrelation mit dieser Dimension von Meßwerten ab, die anderen motivthematischen Kategorien

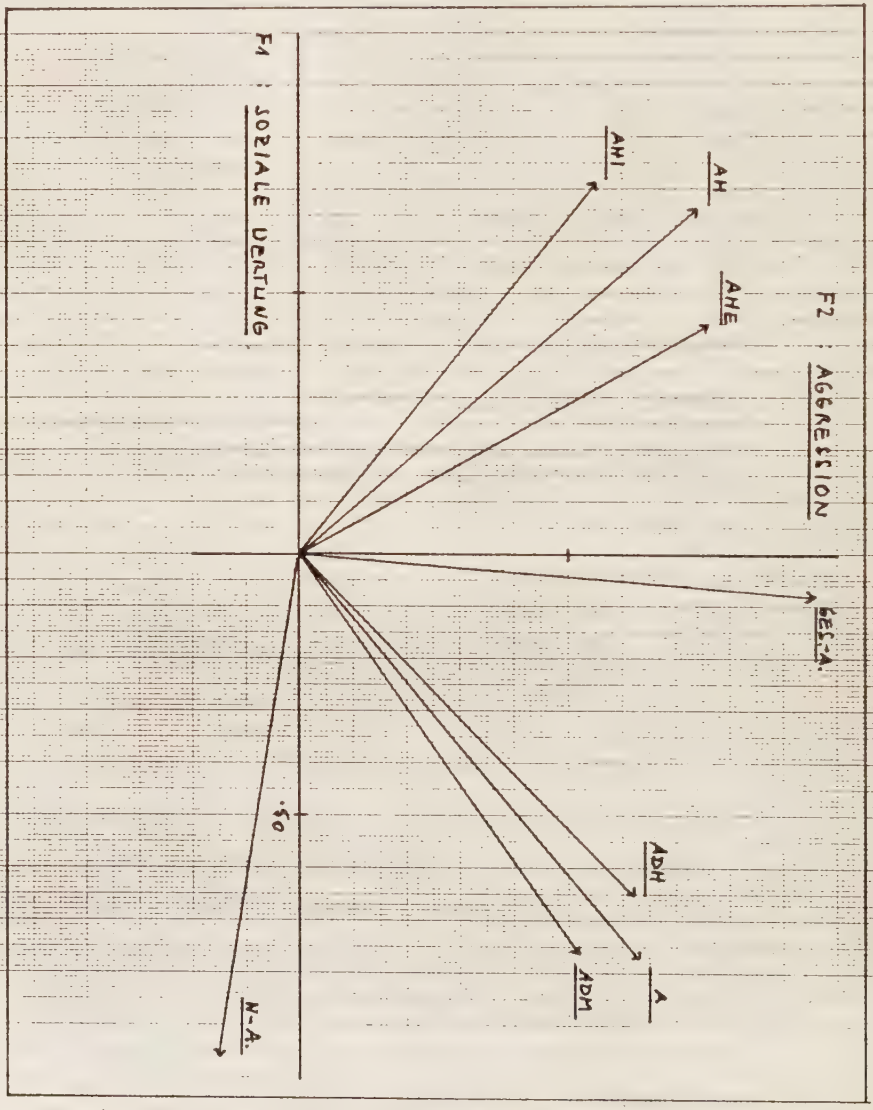


Abb. 14 Vektorgeometrische Darstellung der Dimensionalität der Subskalen des AM-GITTERs im Generalfaktormodell auf Grundlage der in Tab. 12 angegebenen Ladungskoordinaten.

zuzuordnen sind. Mithin handelt es sich bei dieser Dimension um einen motivsystemspezifischen Generalfaktor, den alle fundamentalen Motivkennwerte des Systems laden. Faktor 1 stellt demgegenüber eine Dimension (SOZIALE WERTUNG) dar, mit der die beiden Grundformen aggressiver Motivierung mit unterschiedlichem Vorzeichen korrelieren, da sie einer entgegengesetzten sozialen Wertung unterliegen. Es spricht vorerst nichts dagegen, daß Kennwerte anderer Motivsysteme, die ebenfalls einer sozialen Wertung unterliegen, mit dieser Dimension korreliert sind. Obgleich wir dies noch nicht explizit geprüft haben und dies auch im Rahmen dieser Arbeit nicht vorgesehen ist, könnte diese Dimension einen höheren Allgemeinheitsgrad besitzen und Ausgangspunkt für ein allgemeines Modell zur Dimensionalität sozialer Motive werden (vgl. Abb. 8).

Zu den Ladungsvektoren der abgeleiteten Kennwerte des AM-GITTERs ($\overrightarrow{n\text{-}A.}$, $\overrightarrow{Ges.\text{-}A.}$) äußern wir uns in Abschnitt 2.8.5.

2.8.4 Diskussion

Die nunmehr vorliegenden Ergebnisse sind in mehrfacher Hinsicht von Bedeutung. Zunächst bestätigen sie einmal den von uns eingeschlagenen Weg der Theorieableitung des Verfahrens. Dabei entsprechen die vorliegenden Dimensionen, mit denen wir die Kennwerte des Verfahrens beschreiben, mit ihrem motivsystemübergreifenden, allgemeintheoretischen Charakter gerade unserem theoretischen Ansatz, der darin besteht, daß wir eine allgemeine Theorie der

Motivation auf die thematische Inhaltsklasse Aggression angewandt haben. Diesem zufolge charakterisieren Motivkennwerte auf eine bestimmte Inhaltsklasse bezogene, jedoch (innerhalb eines Motivsystems) unterschiedlich gerichtete Wertungsvoreingenommenheiten. Dabei halten wir es für nicht ausgeschlossen, daß unserer Richtungsdimension eine für "pro-" und "antisoziale Motive" (vgl. HECKHAUSEN 1980, S. 324 ff.) grundlegende Bedeutung beigemessen werden kann.

Zusammenfassend zeichnet sich das nun vorliegende Modell zur Beschreibung der Dimensionalität der Skalen durch folgende Eigenschaften aus:

- hoher Allgemeinheitsgrad der konstituierenden Begriffe (s.o.);
- Einfachheit (niedrige Faktorenzahl);
- Vollständigkeit (Erklärung von 89 % der gesamten Subtestvarianz);
- absolute Objektivität (keine Rotation erforderlich);
- Fruchtbarkeit (gestattet die Ableitung theoretischer Annahmen).

Es braucht nicht eigens hervorgehoben zu werden, daß damit eine ganze Reihe sonst Üblicher Probleme, wie sie sich im Zusammenhang mit der Faktorisierung von Meßwerten stellen (vgl. ÜBERLA 1977), entfällt. Auf einen konkreten Vorteil des letztgenannten Aspektes (Fruchtbarkeit) kommen wir im nächsten Abschnitt zu sprechen. Die Ergebnisse legen es nämlich nahe, die Dimensionalität des Verfahrens theoretisch festzulegen und von dieser Vorgabe die Interkorrelationsstruktur der Subskalen abzuleiten. Damit wollen wir die Konstruktion des AM-GITTERs vorerst abschließen.

2.8.5 Theoretische Festlegung der Dimensionalität der STD-Skalen des AM-GITTERs

2.8.5.1 Voraussetzungen

Die theoretische Festlegung der Dimensionalität von Subskalen scheint uns unter bestimmten Voraussetzungen als durchführbar, auf die wir zunächst kurz eingehen möchten.

Die erste, allgemeinere Voraussetzung besteht u.E. einfach darin, über ein Modell zu verfügen, für das empirische Hinweise darauf bestehen, daß es die zwischen den Subtests bestehende Kovariation auch hinreichend erfaßt, da ansonsten die Gültigkeit des Modells im Hinblick auf diese Skalen stark eingeschränkt wäre. Nach der uns vorliegenden Information können wir dieser Voraussetzung insofern positiv entsprechen, als nahezu 90 % der gesamten Subtestkovariation durch das vorliegende Generalfaktormodell erklärt werden können.

Die theoretische Festlegung der Dimensionalität der Subskalen läuft nun speziell darauf hinaus, für die Vektoren der fundamental messenden Subtests eine bestimmte Zeigerstellung im Modell festzuschreiben. Das Ausmaß der differentiellen Determination eines Meßwertes durch die Dimensionen des Modells muß jedoch zuvor theoretisch begründet werden können, damit es überhaupt (und damit die o.g. Zeigerstellung) vorgegeben werden kann.

Wir verfügen im Augenblick nur über hinreichende Information, um diese zweite, speziellere Voraussetzung für die beiden fundamental

messenden STD-Skalen (A, AH) zu formulieren. Zudem haben wir dort auch schon während der Konstruktion systematisch auf einen bestimmten Winkel phi zwischen deren Testvektoren geachtet, der gerade 90 Grad betragen sollte (10).

Motivationstheoretisch gibt es nun keinen Grund anzunehmen, daß zwischen den beiden Skalen unterschiedliche Gewichtungsverhältnisse im Hinblick auf die beiden Dimensionen des Generalfaktormodells bestehen oder daß beide durch die o.g. Dimensionen in einem je unterschiedlichen Ausmaß beschrieben werden sollten. Es ist vielmehr plausibel - und entspricht geradezu, wie wir noch sehen werden, dem KORNADTschen Strukturmodell des Aggressionsmotiv-Systems - anzunehmen, daß beide Dimensionen des Generalfaktormodells gleichermaßen zur Determination einer fundamental messenden STD-Skala beitragen sollten. Das heißt nun genauer formuliert:

Im Generalfaktormodell zeichnet sich ein Test zur Messung des überdauernden Aggressionsmotivs dadurch aus, daß er zu 50 % durch die thematische Dimension AGGRESSION und zu weiteren 50 % durch die Dimension SOZIALE WERTUNG bestimmt wird. Gleiches gilt für einen Test zur Messung der überdauernden Aggressions-Hemmung. Da jedoch beide nicht dasselbe messen und ihre Meßgegenstände zudem einer entgegengesetzten sozialen Wertung unterliegen, ist es erforderlich festzulegen, daß sie die Richtungsdimension SOZIALE WERTUNG mit unterschiedlichem Vorzeichen laden.

Dadurch erreicht man aber nun gerade eine Kollinearität der Subtestvektoren $\vec{A}$ und $\vec{AH}$ mit der 1. und 2. Winkelhalbierenden des

Generalfaktormodells und - dadurch bedingt - deren Orthogonalität sensu KORNADT.

Zusammengefaßt ergeben sich somit folgende motivationstheoretischen Postulate zur Einordnung der Subtestvektoren $\vec{A}$ und $\vec{AH}$ in das Generalfaktormodell:

(a) beide Subtests laden die thematische Dimension AGGRESSION;

(b) beide Subtests laden die Richtungsdimension SOZIALE WERTUNG;

(c) die beiden Subtests laden die Dimension SOZIALE WERTUNG entgegengesetzt;

(d) die beiden Subtests sind durch die beiden Dimensionen des Modells jeweils gleichanteilig (zu je 50 %) bestimmt.

Die oben gemachten Voraussetzungen entsprechen der definitorischen Festlegung von Motivsystemkomponenten, die man verkürzt so umschreiben könnte: auf eine bestimmte thematische Kategorie bezogene, jedoch unterschiedlich gerichtete, überdauernde Wertungsvoreingenommenheiten.

2.8.5.2 Vektorielle Einordnung der STD-Skalen in das Generalfaktormodell

Aufgrund der o.g. Voraussetzungen sind wir nun in der Lage, den beiden fundamental messenden STD-Skalen, A und AH, einen ganz bestimmten Platz im Modell zuzuweisen. Wir ordnen daher zunächst die beiden Vektoren $\vec{A}$ und $\vec{AH}$ kollinear zur 1. und 2. Winkelhalbierenden in das Modell ein (vgl. FRUCHTER 1954, S. 33 ff.; THURSTONE 1969, S. 87 ff.). Da die Meßwerte der beiden abgeleiteten

STD-Skalen (Ges.-A., n-A.) durch Summen-/Differenzenbildung aus den beiden fundamentalen Meßwerten definiert sind, lassen sich die Testvektoren $\overrightarrow{\text{n-A.}}$ und $\overrightarrow{\text{Ges.-A.}}$ nach den Regeln der Vektoraddition einordnen:

$$\vec{A} + \vec{AH} = \overrightarrow{\text{Ges.-A.}}$$

$$\vec{A} - \vec{AH} = \overrightarrow{\text{n-A.}}$$

Wir vollziehen diese Operationen graphisch nach und erhalten die Abbildung 15.

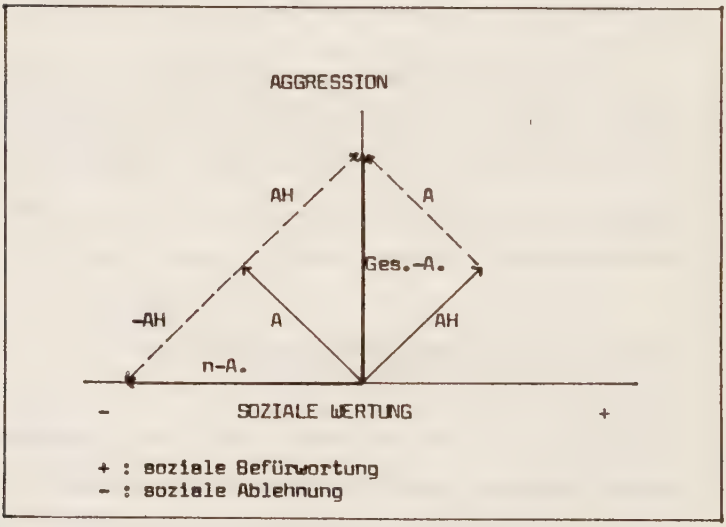


Abb. 15 Vektorielle Einordnung der STD-Skalen in das Generalfaktormodell

Innerhalb dieses Systems (vgl. Abb. 15) läßt sich nun die Korrelation zwischen zwei beliebigen Tests nach Gleichung (1) als Skalarprodukt ihrer Testvektoren bestimmen.

$$(1) \qquad r_{ij} = h_i h_j \cos phi_{ij}$$

Nun normieren wir das System (Abb. 15) auf die Einheitlänge 1 und erhalten damit Abbildung 16.

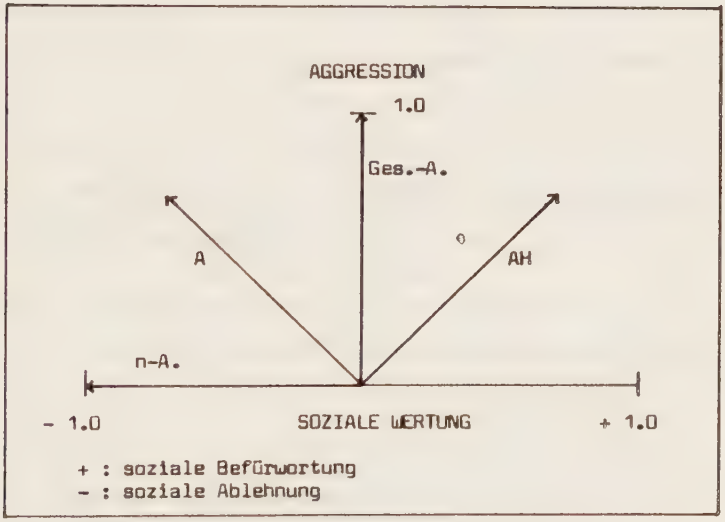


Abb. 16 Vektorielle Einordnung der STD-Skalen in das auf die Einheitslänge 1 normierte Generalfaktormodell

Aufgrund der Normierung des Systems (Abb. 16) vereinfacht sich die Beziehungsgleichung (1) zur Gleichung (2).

$$(2) \qquad r_{ij} = \cos phi_{ij}$$

Nach Gleichung (2) sind nun die Ladungsmatrix (Korrelation der Subtests mit den Dimensionen als Referenzvektoren) als auch die Interkorrelationsmatrix der STD-Skalen durch die jeweils zwischen den Testvektoren bzw. zwischen Test- und Referenzvektoren eingeschlossenen Winkel (phi_{ij}) festgelegt.

Die theoretisch festgelegte Ladungsmatrix der STD-Skalen ist Tabelle 13 zu entnehmen.

	DIMENSION	
SUBTESTVARIABLE	AGGRESSION	SOZIALE WERTUNG
A	$\cos 45^0 = .71^+$	$\cos 135^0 = -.71^+$
AH	$\cos 45^0 = .71^+$	$\cos 45^0 = .71^+$
n-A.	$\cos 90^0 = .00$	$\cos 180^0 = -1.0$
Ges.-A.	$\cos 0^0 = 1.0$	$\cos 90^0 = .00$

+ approximativ

Tab. 13 Theoretisch festgelegte Koordinaten der STD-Skalen des AM-GITTERs im Generalfaktormodell

Wie wir einem Vergleich von Tabelle 12 mit Tabelle 13 entnehmen können, entsprechen die empirisch aufgefundenen Faktorladungen der STD-Skalen im unrotierten Generalfaktormodell den

theoretisch geforderten gut und bestätigen unseren Verzicht auf Nachrotation, mit dem wir uns zudem die absolute Objektivität des Modells erhalten haben. Die Vorzeicheninkongruenzen in den Koordinaten sind hier nicht weiter von Bedeutung; wir haben uns entschieden, den Ast "Soziale Befürwortung" der Dimension SOZIALE WERTUNG nach rechts zu legen, was den normalen Lesegewohnheiten mehr entspricht. Die Rechts-Links-Kongruenz einer empirischen Ladungsmatrix mit der o.g. Vorgabe hängt z.B. davon ab, ob wir weitere Kennwerte (z.B. für FREUNDSCHAFT; vgl. Abb. 6, Abb. 7) mitfaktorisieren. Aus diesem Grunde verbietet es sich auch, die Vorzeichen (+, -) der Dimension SOZIALE WERTUNG phänomenologisch zu interpretieren; sie verdeutlichen lediglich die beiden einander entgegengesetzten Äste der Dimension. In diesem Sinne sind auch die Vorzeichen der in Tabelle 13 dargestellten Koordinaten auf dieser Dimension zu verstehen.

Die Korrelationsstruktur zwischen den STD-Skalen läßt sich auf die gleiche Art und Weise theoretisch vorgeben, wie in Tabelle 14 dargestellt. Sie ist invariant gegenüber einer Rotation des Systems, womit die o.g. Inkongruenzen entfallen. Sie läßt sich daher unproblematisch als (theoretischer) Maßstab zur Beurteilung einer empirisch aufgefundenen verwenden. Aufgrund der geschilderten Zusammenhänge zwischen Interkorrelations- und Ladungsmatrix läßt sich auch die Güte der Übereinstimmung zwischen empirischer und theoretischer Ladungsmatrix durch einen Vergleich der zugehörigen Korrelationsmatrizen feststellen.

r_{xy}	A	AH	n-A.	Ges.-A.
A		$\cos 90^{\circ}$ = .00	$\cos 45^{\circ}$ = .71+	$\cos 45^{\circ}$ = .71+
AH			$\cos 135^{\circ}$ = -.71+	$\cos 45^{\circ}$ = .71+
n-A.				$\cos 90^{\circ}$ = .00
Ges.-A.				

+ approximativ

Tab. 14 Theoretisch festgelegte Interkorrelationsmatrix der STD-Skalen des AM-GITTERs im Generalfaktormodell

Wie wir einem Vergleich von Tabelle 11 mit Tabelle 14 entnehmen können, stimmt die empirisch aufgefundene Korrelationsstruktur zwischen den STD-Skalen des AM-GITTERs gut mit der theoretisch vorgegebenen überein und belegt damit, daß wir unsere konstruktiven Zielsetzungen diesbezüglich als erreicht betrachten können.

2.9 Zusammenfassung

Man kann zusammenfassend sagen, daß sich die von uns gewählte Strategie zur Entwicklung des AM-GITTERs, nach den Voruntersuchungen im großen und ganzen selektiv weiterzuverfahren und ohne größere konstruktive Änderungen in den situations- und personseitigen Bestandteilen des Verfahrens auszukommen, bewährt hat. Wir verfügen nun über ein Instrumentarium, das mit 8 verschiedenen, gut differenzierenden, meßgenauen und dimensionsanalytisch geprüften Skalen verschiedene Aspekte des Aggressionsmotiv-Systems zu erfassen gestattet.

Vier Skalen dienen der Standardauswertung des AM-GITTERs. Zwei davon messen fundamental das Aggressionsmotiv (Skala A) und die Aggressions-Hemmung (Skala AH); die beiden übrigen dienen - in Anlehnung an die Auswertung des A-TAT - zur (abgeleiteten) Messung der Gesamt-Aggressivität (Skala Ges.-A.) und der Netto-Aggression (Skala n-A.). Die Dimensionalität dieser vier Skalen wurde zum Schluß der Hauptuntersuchung theoretisch berechnet. Es besteht dabei eine gute Übereinstimmung zwischen den theoretisch vorgegebenen Koordinaten im (Generalfaktor-) Modell und den empirisch ermittelten.

Vier weitere Skalen des AM-GITTERs beruhen auf dimensionsanalytisch belegten Aspekten der Binnendifferenzierung innerhalb des Aggressionsmotivs einerseits (Skala ADM, Skala ADH) und innerhalb der Aggressions-Hemmung andererseits (Skala AHI, Skala AHE).

Diese differentiellen Skalen sind ein unmittelbares Ergebnis unserer Bemühungen, die Dimensionalität der Meßwerte des AMGITTERs zu berechnen. In diesem Zusammenhang sind wir zu einer veränderten Auffassung über die formale Konzeptualisierung des Aggressionsmotiv-Systems gelangt, ohne dies ursprünglich beabsichtigt zu haben. Vielmehr haben Analysen zur Dimensionalität der Motivierungsdaten wiederholt diese Auffassungsänderung nahegelegt.

Ausgegangen waren wir von einem Strukturmodell zweier Hauptkomponenten, Aggressionsmotiv und Aggressions-Hemmung. Es zeigte sich aber, daß sich nach diesem Modellansatz mehr als zwei Hauptkomponenten in den Daten verselbständigen und sich eine weitere Modellvariante mit vier Hauptkomponenten etablieren läßt. Die neuen Komponenten gingen dabei gewissermaßen aus einer Aufspaltung der beiden o.g. Hauptkomponenten hervor. Die Feinheiten dieser dimensionalen Untergliederung zeigten sich auf Analyse-Ebene von Statements in der Form eigenständiger Dimensionen. Auf Subskalen-Ebene traten sie jedoch in dieser Form wieder zurück, blieben aber dennoch (in anderer Form) erhalten.

Weiterführende Überlegungen zeigten, daß es sinnvoll ist, die Meßdaten in einem zweifaktoriellen Generalfaktormodell zu beschreiben. Wir konnten das auf unterschiedlichem Niveau der Zusammenfassung von Testitems (Statement-Summen, Subskalen-Rohwerte) nachweisen. Damit rückten wir jedoch von tradierten Vorstellungen ab, ohne allerdings zu diesen in einen direkten Widerspruch zu geraten.

Zunächst faßt das Generalfaktormodell zentrale Aspekte beider Hauptkomponentenmodellvarianten integrativ zusammen und dient somit der Vereinfachung der Befundlage. Daneben bietet es eine Reihe formaler Vorteile und einen höheren Grad an Objektivität.

Wie wir gesehen haben, entspricht die varimaxrotierte Zweifaktorenlösung, mit der wir das zweidimensionale Hauptkomponentenmodell des Aggressionsmotiv-Systems stützen, einem um etwa 1/8 Volldrehung nachrotierten Generalfaktormodell. Wenn wir dessen Dimensionen jedoch durch Nachrotation zur Deckung mit den motivationstheoretischen Komponenten bringen, können wir nicht mehr erwarten, daß sie diese damit noch dimensional erklären; vielmehr gehen damit alle Vorteile der unrotierten Zweifaktorenlösung wieder verloren, und wir rekurrieren auf ein weniger leistungsfähiges Modell. Durch die unipolare Dimension des Generalfaktormodells (AGGRESSION) entsprechen wir dem definitorisch festgelegten Bezug sowohl des Aggressionsmotivs als auch der Aggressions-Hemmung auf die thematische Kategorie Aggression, durch die bipolare Dimension (SOZIALE WERTUNG) dem unterschiedlich gerichteten Bezug der o.g. Komponenten hinsichtlich deren sozialer Wertung. Wir haben uns speziell aufgrund der Ladungsvektoren der freundschaftsthematischen Statements zu diesem Begriff entschlossen, da er auf eine Gemeinsamkeit zwischen Verhinderung feindseliger Aggression und Aufrechterhaltung freundschaftlicher Beziehungen verweist.

Einen besonderen Vorteil bietet das Generalfaktormodell hinsichtlich

der meßtheoretischen Begründung abgeleiteter Kennwerte für Netto- und Gesamtmotivation. Diese, in Anlehnung an HECKHAUSEN (1963) vorgenommene Bildung von Motiv-Kennwerten wurde bislang nicht wirklich theoretisch begründet (vgl. z.B. BÄUMLER & WEISS 1967; KUHL 1978). Man konnte davon ausgehen, daß sie - auf grobem Niveau - eine sinnvolle Information liefern. Bei sehr hoher Gesamtmotivation kann man zu dem Schluß eines stark konflikthaft strukturierten Motivsystems gelangen, und das Vorzeichen des Netto-Kennwertes sagt etwas darüber aus, ob es sich um eine vorwiegend aufsuchens- oder meidensmotivierte Person handelt. Eine meßtheoretische Begründung der additiven Verknüpfung unterschiedlich gerichteter Motivkennwerte ist das natürlich nicht, worauf insbesondere BÄUMLER & WEISS (1967) hingewiesen haben. Unser Modell setzt sich jedoch von der von BÄUMLER & WEISS (1967) vertretenen Zweifaktorentheorie der Leistungsmotivation (vgl. auch BÄUMLER & DVORAK 1969; BÄUMLER & BREITENBACH 1970) vor allem dadurch ab, daß es auf Begriffen beruht, die mehr allgemeintheoretischen Charakter und weniger motivsystemspezifischen haben (vgl. ebd.).

Aufgrund der Tatsache, daß unser Generalfaktormodell nicht nur auf Meßwerte-Ebene, sondern gerade auch schon auf Definitions-Ebene des Aggressionsmotiv-Systems Gültigkeit besitzt, gestattet es auch die Ableitung von Annahmen, zu welchen Anteilen je eine Motivkomponente durch seine Dimensionen erklärt bzw. festgelegt werden soll. Von einer Vorstellung der jeweils gleichanteiligen Determination des Aggressionsmotivs als auch der Aggressions-Hemmung durch die beiden Dimensionen ausgehend läßt sich insbesondere

die im KORNADTschen Zweikomponentenmodell geforderte Orthogonalität der Struktur theoretisch begründen. Darauf aufbauend konnten wir anschließend Interkorrelations- und Ladungsmatrix der STD-Skalen theoretisch berechnen.

Innerhalb des Generalfaktormodells zeichnen sich die 8 Subskalen, die wir entwickelt haben, durch eine unterschiedliche Zeigerstellung ihrer Testvektoren aus. Die Zeigerstellung impliziert eine differentielle Gewichtung in der Determination der Meßwerte einer Skala durch die beiden Dimensionen des Modells. Während die beiden fundamentalen STD-Skalen jeweils gleichanteilig zu je 50 % durch die eine und zu den restlichen 50 % durch die andere der beiden Dimensionen bestimmt werden, ist die abgeleitete STD-Skala n-A. zu 100 % durch die Dimension SOZIALE WERTUNG und die abgeleitete STD-Skala Ges.-A. zu 100 % durch die thematische Dimension AGGRESSION bestimmt. Durch Differenzenbildung maximieren wir also das Unterscheidende, durch Summenbildung das Gemeinsame zwischen den beiden fundamentalen Meßwerten in einem neuen. Wir gehen dabei, das sei in diesem Zusammenhang insbesondere hinsichtlich des letzteren (Ges.-A.) erwähnt, zunächst noch von einer neutralen Mitte des bipolaren Faktors aus.

Wenn man die beiden fundamental messenden STD-Skalen in das Generalfaktormodell eingeordnet hat, zerlegen sie selbiges in 4 Sektoren, die - im theoretisch idealen Fall - den jeweiligen Hälften des 1. und 2. Quadranten entsprechen. Die DIFF-Skalen zeichnen

sich nun dadurch aus, daß sie - nach links und rechts von den fundamentalen STD-Skalen abgerückt - in diesen Sektoren liegen. Damit bieten sie zumindest formal die Möglichkeit, durch Messung etwas über die "innere Dynamik" des Feldes in Erfahrung zu bringen, womit sie wesentlich zu einer weiterführenden Interpretation des Modells beitragen könnten. Im theoretisch idealen Fall wiederum hätte man sie sich um etwa 22.5 Grad von den fundamentalen STD-Skalen abgerückt und im System von gleicher Kommunalität (Länge der Testvektoren) vorzustellen. Unter diesen Voraussetzungen hätte man das Feld, zusammen mit den abgeleiteten STD-Skalen, gleichabständig durch Subtestvektoren unterteilt. Man muß dabei jedoch beachten, daß man eine abgeleitete Skala nicht unmittelbar mit einer kollinearen, gleichgerichteten Fundamentalskala gleichsetzen kann, zumal dann, wenn sich eine Linearkombination aus Meßwertebestandteilen zusammensetzt, die zum Teil im ersten, zum Teil im zweiten Quadranten liegen, wie das bei den abgeleiteten STD-Skalen der Fall ist. Wenn dem nicht so wäre, könnte man generell die beiden fundamentalen STD-Skalen (allgemein die beiden in unterschiedlichen Quadranten liegenden Fundamentalskalen, die die längsten Testvektoren aufweisen) als Referenzvektoren betrachten und jede beliebige Stelle des Feldes durch eine spezielle Linearkombination erreichen. In diesem Fall müßte man die DIFF-Skalen von vornherein als redundant betrachten. Sie sind es aber (aller Voraussicht) nicht, denn die Diskriminationsfähigkeit einer wie oben beschriebenen Linearkombination ist natürlich nur in dem Fall gegeben, daß ihre beiden Einzelkomponenten

gleichzeitig Diskriminationsrelevanz besitzen. Davon kann man jedoch nicht generell ausgehen, und es ist vorerst zudem noch völlig unklar, unter welchen Bedingungen man sie im AM-GITTER überhaupt voraussetzen kann.

Die 8 Skalen, die wir nun entwickelt und vorgestellt haben, lassen - bei der relativen Kürze des Verfahrens - schon eine recht differenzierte Auswertung der mit dem AM-GITTER erhobenen Motivierungsdaten zu. Nichtsdestoweniger harren alle Meßwerte einer weiterführenden Aufklärung ihrer inhaltlichen Bedeutung, wozu erste Schritte im nun folgenden Validierungsteil dieser Arbeit unternommen werden sollen.

3. Untersuchungen zur Validität des AM-GITTERs

3.1 Einleitende Gesichtspunkte

Mit der Entscheidung für eine Theorieableitung beginnt die Validierung eines Verfahrens mit den ersten Schritten seiner Konstruktion. Daher sind die Iteminhalte unseres GITTERs bereits auf phänomenologischer Ebene unmittelbar einsichtig und bescheinigen ihm logische Validität für den beanspruchten Gültigkeitsbereich. Die ermittelten Meßwerteverteilungen auf den einzelnen Skalen, die Bestimmung deren Reliabilität zum Zeitpunkt der Messung als auch die wiederholt durchgeführten dimensionsanalytischen Untersuchungen, die wir konstruktiv in die Entwicklung unseres Instruments einbezogen haben, entsprechen im weiteren Forderungen, wie sie etwa von LIENERT (1969, S. 261 ff.) in Anlehnung an die von CRONBACH & MEEHL entworfenen Vorstellungen zur Konstruktvalidierung erhoben werden (vgl. S. 262, Punkte 5 und 7, ebd.).

Was die Eigenständigkeit dieses Teils der Arbeit begründet ist, daß wir nunmehr Schritte einleiten, um die Gültigkeit unseres Verfahrens unter weiterführenden Gesichtspunkten zu prüfen, ohne daß dabei - wie bislang - die Prüfergebnisse unmittelbar konstruktive Veränderungen zur Folge haben.

Ziel der nachfolgend durchgeführten Untersuchungen ist es, die Validität des AM-GITTERs in Ergänzung zu den bereits vorliegenden Hinweisen darauf zunächst durch folgende Einzelmaßnahmen

weiter zu prüfen:

- Analyse der Veränderung individueller Kennwerte bei wiederholter Durchführung der Messung ohne systematische Variation der Durchführungsbedingungen;
- Analyse der Veränderung Gruppen-bezogener Kennwerte bei wiederholter Durchführung der Messung ohne systematische Variation der Durchführungsbedingungen;
- Korrelation der Meßwerte mit Meßwerten ähnlichen Validitätsanspruchs;
- Korrelation der Meßwerte mit Meßwerten, die andere Persönlichkeitsmerkmale erfassen;
- Korrelation der Meßwerte mit Außenkriterien;
- Faktorisierung der Meßwerte gemeinsam mit Außenkriterien

(vgl. dazu LIENERT 1969, S. 262, Punkte 1 bis 4 und Punkt 6).

Aus didaktischen Gründen berichten wir zunächst über diejenigen Analysen, die mehr auf die formalen Aspekte der Validität Bezug nehmen. Wir untersuchen dabei das Verhalten der Meßwerte des AM-GITTERs bei Wiederholungsmessungen nach unterschiedlich langen Zeitintervallen. Sodann stellen wir die Analysen vor, die mehr auf die inhaltlichen Aspekte der Validität eingehen. Dazu gehören Berechnungen zur "diskriminanten" und "konvergenten Validität" sowie Zusammenhangsanalysen zwischen Meßwerten des AM-GITTERs und verschiedenen Außenkriterien. Letztere ergänzen wir um eine gemeinsame Faktorisierung der Meßwerte beider Erhebungsmodalitäten.
Zum Schluß berichten wir über eine gesondert (ex post facto) durchgeführte Analyse, die zum Ziel hat, einen Detailbefund theoretisch

weiter aufzuhellen. Wir gehen in dieser Analyse Zusammenhängen zwischen Aggressivität und Attribuierungsverhalten nach.

3.2 Meßwertestabilität

3.2.1 Vorüberlegungen

Abgesehen von empirischen Befunden, die darauf hinweisen, daß es sich bei der menschlichen Aggressivität um ein zeitlich relativ stabiles Persönlichkeitscharakteristikum handelt (LEFKOWITZ et al. 1977; OLWEUS 1979, 1980; vgl. auch PARKE & SLABY 1983), stellt man sich Motive im allgemeinen (McCLELLAND 1958; HECKHAUSEN 1980) und so auch das Aggressionsmotiv-System im besonderen (KORNADT 1982 a) als relativ zeitstabile Dispositionsvariablen der Persönlichkeit vor, obgleich unsystematisch oder systematisch induzierte Veränderungen in der Stärke der Motivausprägung grundsätzlich eingeräumt bleiben (vgl. z.B. KRUG 1976; KRUG & HANEL 1976). Insofern sollte ein Meßwert, der bestimmte Aspekte eines Motivs zu erfassen beansprucht, dieses formale Charakteristikum ebenfalls aufweisen.

McCLELLAND (1958, S. 18 ff.) hat das Stabilitätspostulat für Motivmeßwerte mit seinem dritten Kriterium für ebensolche in folgende Worte gefaßt: "The measure of a motive should give the same reading for an individual or a group under the same or nearly the same conditions."

Bezogen auf Individuen wird die Meßwertestabilität üblicherweise

als Retestkorrelation ermittelt. Aus einer Vielzahl theoretischer Erwägungen heraus läßt sich jedoch kaum begründen, wie hoch diese ausfallen soll (DIETERICH 1973, S. 148 ff.; RAUCHFLEISCH 1980, S. 52 f.), speziell für Motivmeßwerte (HALISCH 1982, S. 149 f.; McCLELLAND 1980, S. 29 ff.; 1981, S. 97 f.).

Die Gründe dafür sind z.B. darin zu suchen, daß i.d. Regel geeignete Möglichkeiten fehlen, exakt unter den gleichen Bedingungen wiederholt zu messen bzw. den Gütegrad abzuschätzen, mit dem die Meßbedingungen replizierbar waren. Dieser Faktor dürfte umso mehr zu Buche schlagen, je genauer ein Motivtest mißt und je mehr die Meßsituation motivanregungsrelevante Hinweisreize beinhaltet.

Ein noch gänzlich unberechenbarer Faktor sind azyklisch oder zyklisch stattfindende Merkmalsfluktuationen auf der Grundlage kognitiv-physiologischer Wechselwirkungen, die sich (nach der o.g. Bestimmungsmethode) reliabilitätsmindernd auswirken können (vgl. RAUCHFLEISCH 1980, S. 53).

Man wird jedoch mit einigem Recht die Annahme vertreten können, daß eine Wiederholungsmessung, nach verschieden langen Zeitintervallen aufgenommen, zumindest überzufällig mit der Erstmessung korreliert sein sollte. Darin kommt die Vorstellung zum Ausdruck, daß das über diesen Zeitraum stabile Motiv vorhersagbar die Höhe beeinflußt, in der die Messung ausfällt. Zwar kann man bei nicht nachgewiesener Stabilität nicht notwendig auf eine unzuverlässige

Messung schließen, wie DIETERICH (1973, S. 148) meint - zumindest solange keine exakte Kontrolle der Meßbedingungen sowie der zwischenzeitlich stattgefundenen Prozesse möglich ist -, jedoch kann man umgekehrt bei nachgewiesener Stabilität diese sehr wohl als eine konservative (untere) Schätzung der Meßgenauigkeit eines Verfahrens (unter den angegebenen empirischen Bedingungen) betrachten.

Im Hinblick auf die eingangs genannten Aspekte der Konstruktvalidität untersuchen wir im folgenden zunächst intraindividuelle Veränderungen bei wiederholter Durchführung der Messung ohne systematische Variation der Durchführungsbedingungen. Sodann betrachten wir Veränderungen in den Meßwerteverteilungen zwischen Erst- und Zweitmessung. Im Hinblick auf das zitierte McCLELLAND-Kriterium beziehen wir uns also zunächst auf Individuen und sodann auf Gruppen.

Wir erhoffen uns von diesen Maßnahmen erste ergänzende Informationen zu den bereits aus den Konsistenzanalysen und anderen Berechnungsansätzen vorliegenden Reliabilitätsbefunden.

3.2.2 Ausgangshypothesen

Wir untersuchen nun das Verhalten der Meßwerte des AM-GITTERs bei Wiederholungsmessungen nach unterschiedlich langen Zeitintervallen. Wir erwarten dabei unabhängig von der Größe des Zeitintervalls zwischen Erst- und Zweitmessung

(a) signifikante Retestkorrelationen für alle Skalen des AM-GITTERs;

(b) Gleichheit der Verteilungsfunktionen aus Erst- und Zweitmessung für alle Skalen des AM-GITTERs.

3.2.3 Methode

Aus der Eichstichprobe (vgl. Abschnitt 2.3.1) wählen wir vier Grundschulen zum Retest aus. Untersucht werden alle dort eingerichteten 4. Klassen im Gruppenversuch, an dem auch die Mädchen teilnehmen, insgesamt 10 Klassen. Die Datenaufnahme erfolgt im Zeitraum November 1980 bis Januar 1981. Die Testungen werden morgens während der normalen Unterrichtszeit durchgeführt gemäß den Bedingungen in der Hauptuntersuchung; der Testleiter ist jeweils der gleiche, besondere Vorkommnisse während der Testungen sind nicht zu vermerken.

Zur Einführung wird den Kindern, die nicht darauf vorbereitet sind, das Verfahren erneut zu bearbeiten, mitgeteilt, daß die Ergebnisse ihrer Klasse (im Erstversuch) so überraschend gut ausgefallen seien, daß man es kaum glauben könne, wenn es sich nicht noch einmal bestätigen lasse. Verdeutlicht wird das den Kindern am Beispiel eines Jungen, der behauptet, eine bestimmte Höhe im Hochsprung zu erreichen und das vor seinen Freunden "beweisen" muß. Auf diese Weise kann bei den meisten Kindern ein ersichtlicher Leistungsstolz induziert werden, zumal ihnen anschließend auch noch mitgeteilt wird, daß die übrigen Klassen weit schlechter abgeschnitten haben und sie gerade deshalb noch einmal erneut

am Versuch teilnehmen sollen. Danach darf sich jedes Kind ein Testheft holen. Kindern, die nicht teilnehmen möchten, wird die Möglichkeit geboten, sich währenddessen still zu beschäftigen.

Durch die o.g. Maßnahme soll ausgeschlossen werden, daß Kinder die schon bekannte Aufgabenstellung mit Widerwillen bewältigen, d.h. Frustrationseffekte auftreten. Durch eine ggfls. durch die situative Einbettung der Retestung bedingte mitinduzierte, andauernde Test(leistungs-)motivation sollten die Meßwerte ohnehin nicht beeinflußt werden; dies entspräche - falls dem so wäre - zudem einer systematischen Änderung der Durchführungsbedingungen und sollte sich eher reliabilitätsmindernd auswirken. Wir haben dies jedoch nicht weiter geprüft.

Als Zeitintervalle für die Durchführung der Wiederholungsmessungen wählen wir 2, 4 und 8 Wochen.

Im einzelnen resultieren danach folgende Beteiligungszahlen an Jungen, deren Daten in die Endverrechnung eingehen (vgl. Tab. 15):

RETESTINTERVALL	ANZAHL JUNGEN
2 Wochen	12
4 Wochen	27
8 Wochen	26

Tab. 15 Gesamtstichprobe und Aufteilung der Probanden bei den Wiederholungsmessungen

Wir bestimmen die Stabilität der Motivkennwerte als Rangkorrelation zwischen den Rohwerten aus Erst- und Zweitmessung. Die Signifikanz der Koeffizienten prüfen wir mit dem t-Test von FISHER (CLAUSS & EBNER 1975, S. 274 f.), zweiseitig. Das gewählte Programm führt zusätzlich eine Bindungskorrektur durch.

Ein Motivtest, wie wir ihn zu erstellen beabsichtigen, dient nicht nur der Charakterisierung von Einzelpersonen (idiographischer Verwendungszweck), sondern vor allem auch der Beschreibung von Gruppen (nomothetischer Verwendungszweck). Auf dieser Ebene war zunächst an einen Einsatz unseres Verfahrens als theoriebezogenes Meßinstrument gedacht.

Man charakterisiert die Meßwerteverteilung einer Gruppe von Individuen i.d. Regel durch Angabe ihres arithmetischen Mittelwertes sowie der zugehörigen Varianz. Stabilität einer Meßwerteverteilung bedeutet, daß bei Meßwiederholung in diesen Parametern keine überzufällige (systematische) Veränderung eintritt. Die vorliegenden Motivkennwerte lassen eine Prüfung dieser Annahme zu. Die Nullhypothese, Gleichheit der Verteilungsfunktionen aus der Test- bzw. Retestphase ($F_1(x) = F_2(x)$), prüfen wir mit dem RANG-VORZEICHEN-TEST von WILCOXON (SACHS 1972, S. 244 ff.). Die Prüfstatistik reagiert sensibel auf Mittelwerteveränderungen bei Varianzhomogenität bzw. auf kombinierte Mittelwerte-Varianz-Effekte.

3.2.4 Ergebnisse

3.2.4.1 Stabilität individueller Kennwerte

Die Ergebnisse sind im einzelnen der Tabelle 16 zu entnehmen.

RANGKORRELATION		M E S S Z E I T I N T E R V A L L		
AM-Gitter-SKALA		2 WOCHEN	4 WOCHEN	8 WOCHEN
STD	A	.82 +++	.71 +++	.75 +++
	AH	.57 +	.45 +	.77 +++
	n-A.	.75 ++	.56 ++	.64 +++
	Ges.-A.	.73 ++	.67 +++	.70 +++
DIFF	ADM	.74 ++	.63 +++	.62 +++
	ADH	.79 ++	.62 +++	.72 +++
	AHI	.46 n.s.	.58 +++	.72 +++
	AHE	.60 +	.38 +	.55 ++
n		12	27	26
df		10	25	24

n.s. nicht signifikant
\+ Signifikanzniveau 5 %
++ " 1 %
+++ " 0.1 %

Tab. 16 Stabilität individueller Motivkennwerte für die Skalen des AM-GITTERs bei unterschiedlichen Meßzeitintervallen
(SI/SPSS, NONPAR CORR)

Wir ermitteln bis auf eine Ausnahme (Wiederholungsmessung mit der DIFF-Skala AHI im Abstand von 2 Wochen) durchgehend signifikante Retestkoeffizienten. Nach einem Intervall von 2 Wochen sind die Meßwerte auf einer, nach einem Intervall von 4 Wochen auf fünf und nach einem Meßzeitintervall von 8 Wochen auf sieben der acht untersuchten Skalen hochsignifikant (1 Promille-Niveau) korreliert.

Naturgemäß konzentriert sich nun unser Interesse auf die einzige nichtsignifikante Korrelation (Meßwerte der DIFF-Skala AHI und 2 Wochen-Intervall). Wir versuchen eine differentielle Berechnung der Retestkorrelation nach Unterteilung der Stichprobe in Hoch- und Niedrigmotivierte. Dazu berechnen wir für jede Versuchsperson aus Erst- und Zweitmessung einen durchschnittlichen Motivkennwert (arithmetisches Mittel) auf der Skala AHI und stellen anschließend fest, ob dieser unter oder über dem für die DIFF-Skala AHI errechneten Median (Md = 14.14; vgl. ANHANG IV) liegt. Auf diese Weise können wir die Stichprobe in zwei Hälften mit jeweils n = 6 als Besetzungszahl splitten. Für die Hochmotivierten errechnen wir nun eine Retestkorrelation von rho = .64 (n.s., zw., t-Test nach FISHER) und für die Niedrigmotivierten eine Retestkorrelation von rho = -.81 (sig. 5 %, zw., t-Test nach FISHER). Während sich die Hochmotivierten also annähernd erwartungstreu verhalten, registrieren wir bei den Niedrigmotivierten einen signifikanten, gegenteiligen Effekt. Es zeigt sich, daß sich in der Korrelation von .46 (s.o.) zwei

entgegengerichtete Tendenzen überlagern. Die Verteilung der Motivkennwerte verändert sich allerdings über dieses Meßzeitintervall hinweg nicht (vgl. Tab. 18, Abschnitt 3.2.4.2).

3.2.4.2 Stabilität der Meßwerteverteilungen

Mittelwerte und Varianzen auf den einzelnen Skalen - getrennt für Erst- und Zweitmessung nach jeweils unterschiedlichen Meßzeitintervallen dargestellt - sind der Tabelle 17 zu entnehmen.

Das Ergebnis der Signifikanzprüfungen ist in Tabelle 18 dargestellt. Unseren Befunden zufolge weisen wir die Nullhypothese (Gleichheit der Verteilungsfunktionen) in 4 Fällen bei den STD-Skalen und in 2 Fällen bei den DIFF-Skalen zurück.

Insbesondere die DIFF-Skalen ADM und AHE zeichnen sich über alle untersuchten Zeitintervalle hinweg durch eine hohe Konstanz ihrer Meßwerteverteilungen aus. Bei allen anderen Skalen verzeichen wir je einmal eine signifikante Veränderung in der Meßwerteverteilung entweder nach einem Zeitraum von 2 oder von 4 Wochen. Nach einem Zeitraum von 8 Wochen sind die Meßwerteverteilungen auf allen Skalen des AM-GITTERs stabil.

Damit können wir das insgesamt günstige Ergebnis aus der Berechnung der Retestkorrelationen nach einem Meßzeitintervall von 8 Wochen nun auch auf der Ebene der Stabilität von Meßwerteverteilungen stützen. Die Meßwerte des AM-GITTERs sind nach diesem Zeitraum Individuen- sowie Gruppen-bezogen stabil.

arith. Mittel / SKALA	MESSZEITINTERVALL 2 WOCHEN TEST	2 WOCHEN RETEST	4 WOCHEN TEST	4 WOCHEN RETEST	8 WOCHEN TEST	8 WOCHEN RETEST
A	29.9	37.2	28.4	29.0	30.5	28.3
AH	40.3	37.0	37.5	29.8	34.6	35.3
n-A.	-10.3	0.2	-9.1	-0.7	-4.1	-7.0
Ges.-A.	70.2	74.2	65.9	58.8	65.2	63.6
ADM	12.1	14.4	11.6	11.8	13.2	12.0
ADH	9.7	13.3	9.9	10.3	9.6	9.6
AHI	15.3	14.3	14.1	11.1	12.8	12.8
AHE	15.9	14.8	15.2	11.6	14.2	14.2
Varianz						
A	200.6	218.2	179.3	160.1	195.9	263.5
AH	183.5	227.2	287.6	284.7	324.2	308.5
n-A.	251.7	432.2	481.6	349.6	516.7	635.9
Ges.-A.	516.5	459.8	452.4	540.1	523.5	508.2
ADM	39.4	39.2	30.3	26.8	51.7	51.8
ADH	29.2	35.5	33.3	30.1	26.4	31.4
AHI	30.1	32.6	50.0	54.4	64.6	70.5
AHE	46.6	52.5	56.6	50.8	61.4	65.7

Tab. 17 Mittelwerte und Varianzen für die Skalen des AM-GITTERs bei unterschiedlichen Meßzeitintervallen (SI/SPSS, FREQUENCIES)

MESSZEIT-INTERVALL	SKALA A	AH	n-A.	Ges.-A.	ADM	ADH	AHI	AHE
2 WOCHEN	02	ns	04	ns	ns	03	ns	ns
4 WOCHEN	ns	02	ns	03	ns	ns	02	ns
8 WOCHEN	ns	ns	ns	ns	ns	ns	ns	ns

Tab. 18 WILCOXON-Signifikanzniveaus (zweiseitig) in % beim Vergleich der Meßwerteverteilungen der Skalen des AM-GITTERs bei unterschiedlichen Meßzeitintervallen (SI/SPSS, NPAR TESTS (WILCOXON))

3.2.5 Diskussion

Wenn wir als ein akzeptables Kriterium die Vorstellung formulieren, daß zwischen Erst- und Zweitmessung nicht unter 50 % der Varianz systematischer Natur sein sollten, läuft das auf die Annahme einer Korrelation von .71 hinaus, die für eine Skala bei drei Wiederholungsmessungen zumindest einmal nachgewiesen werden sollte. Gemessen an diesem etwas strengeren Maßstab (als dem der lediglich nachgewiesenen Signifikanz der Korrelation zwischen Erst- und Zweitmessung) stellen wir fest, daß alle Skalen - bis auf die DIFF-Skala AHE - dieses Soll erfüllen. Die Skala für das Aggressionsmotiv (A) erreicht/überschreitet dieses Kriterium sogar bei jedem der drei Meßzeitintervalle. Im Gegensatz dazu unterliegen diejenigen Anteile der Aggressions-Hemmung, die wir mit der Skala AHE erfassen, möglicherweise in stärkerem Maße situativen Anregungsbedingungen und/oder haben eine stärkere Eigenfluktuation als die mehr intrinsischen Anteile der Aggressions-Hemmung (DIFF-Skala AHI), denn von den Konsistenzschätzungen her (KR 20) sind beide Skalen praktisch gleich meßgenau (KR 20 = .85 für AHE und .86 für AHI; vgl. Tab. 10, Abschnitt 2.7). Es ist aber auch ein anderer Zusammenhang denkbar.

Wenn wir von der in Tabelle 12 Abschnitt 2.8.3 dargestellten Ladungsmatrix des AM-GITTERs ausgehen, zeichnet sich die Skala mit den günstigsten Retestreliabilitäten (A) vor derjenigen mit den ungünstigsten (AHE) dadurch aus, daß erstere im Modell über den

längsten (Kommunalität = .99), letztere über den kürzesten Testvektor (Kommunalität = .77) verfügt; d.h. erstere erfaßt zu 22 % mehr Anteile des Systems, von denen wir annehmen, daß es sich bei ihnen um die überdauernden, motivsystemspezifischen handelt. In der Tat ist auch die Kommunalität der Skala AHI (h^2 = .82) etwas höher als die der Skala AHE (h^2 = .77). Den recht eigenartigen Befund mit der DIFF-Skala AHI und einem Meßzeitintervall von 2 Wochen interpretieren wir vorerst nicht; diese Wiederholungsmessung sollte (an einer etwas größeren Stichprobe) erneut durchgeführt werden, zumal es u.E. einen Grund gibt, dies überhaupt mit der gesamten Untersuchung zur Meßwertestabilität zu tun. Auf diesen Grund wollen wir nun näher zu sprechen kommen.

Bis auf die DIFF-Skala ADM (und vorerst begrenzt AHI) verzeichnen wir bei allen Skalen des AM-GITTERs bei einem Meßzeitintervall von 4 Wochen einen leichten "Knick", d.h.: gemessen an dem 2 Wochen-Intervall fallen die Korrelationskoeffizienten ab und steigen dann nach 8 Wochen wieder an. Dieser "4 Wochen-Knick", von dem wir einmal annehmen wollen, es sei ein Effekt, scheint in unseren Motivierungsdaten nicht ganz zufälliger Natur zu sein. SCHMALT (1974, S. 238 f.) berichtet ihn (allerdings nicht als "Effekt" in unserem Verständnis) ebenso im Zusammenhang mit Wiederholungsmessungen (gleiche Zeitintervalle) durch das Leistungsmotiv-Gitter für Kinder, wo er sich in ganz auffälliger Weise auf allen 4 grundlegenden Kennwerten für Leistungsmotivation in nahezu allen Situationsbereichen zeigt (1 Ausnahme bei 24 situationsbezogenen Retestprofilen der Meßwerte; ebd., S. 239, Tab. 51). Der Knick zeigt sich

auch deutlich für die Situationskomplexe des LM-GITTERs (ebd., S. 240, Tab. 52). Die bei SCHMALT aufgeführte Vermutung über das Zustandekommen dieses Effektes trifft bei uns nicht zu. Disziplinierungsschwierigkeiten bei der Wiederholungsmessung nach 4 Wochen waren in unserem Fall, ebensowenig wie Unlust, nicht zu beobachten, eher das Gegenteil (vgl. Abschnitt 3.2.3). Gegen die Annahme einer ggfls. reduzierten Testmotivation nach einem Zeitintervall von 4 Wochen spricht auch die gänzlich andersartige (konventionelle) Entwicklung der Stabilitätskoeffizienten, die wir probeweise für einen über die freundschaftsthematischen Statements ermittelten, vergleichbaren Kennwert im AM-GITTER berechnet haben (11). Dieser Detailbefund legt die Vermutung nahe, daß es sich bei dem o.g. Effekt um einen allgemein motivspezifischen in der vorliegenden Population (die auch bei SCHMALT der unseren vergleichbar ist) handeln könnte. Jedenfalls besteht u.E. Grund zu der Annahme, daß die Dauer des Zeitintervalls bei einer Wiederholungsmessung zu der ermittelten Höhe der Retestkorrelation in einem überzufälligen Zusammenhang steht, der erst noch enträtselt werden muß. Das Auslassen des 4 Wochen-Intervalls bei der Angabe von Retestreliabilitäten - wie beim LM-GITTER (SCHMALT 1976 a, S. 122 f.; 1976 b, S. 18 ff.) - ist also im vorliegenden Fall nicht begründbar. Unabhängig davon läßt sich feststellen, daß praktisch alle Skalen bei einer Wiederholungsmessung nach 8 Wochen einem "Grenzwert" von etwa rho = .72 zuzustreben scheinen; bemerkenswert ist auch die bei diesem Intervall aufgezeigte Stabilität nun aller

Meßwerteverteilungen (vgl. Tab. 18).

Bei LIENERT (1969, S. 215 f.) etwa finden wir das Argument, daß Retestreliabilitäten unter gewissen Umständen aufgrund von Erinnerungs-/Gedächtniseffekten eine Meßgenauigkeit vortäuschen, die in Wirklichkeit garnicht vorhanden ist. Abgesehen davon, daß, falls ein solcher Effekt auftritt, er kaum zu einer Überschätzung der Zuverlässigkeit beiträgt, sondern eher im Gegenteil die Korrelation mindert, falls er unsystematisch auftritt (DIETERICH 1973, S. 148), können wir für unser Verfahren zunächst einmal von der Argumentation SCHMALTs (1976 b, S. 21) ausgehen, der im Hinblick auf das vergleichbare LM-GITTER äußert, daß es aufgrund seiner Komplexität Erinnerungseinflüsse kaum zulassen dürfte. Daneben treffen auch Argumente, wie sie etwa von McCLELLAND (1980) für das Zustandekommen solcher Einflüsse vorgetragen werden (z.B. einen guten Eindruck machen), für unser Verfahren nicht zu. Hauptsächlich aber glauben wir, daß die durchgehend hohen Retestkorrelationen für den Motivkennwert A als auch die nach dem 4 Wochen-Knick wieder steigenden Koeffizienten gegen die Annahme des Einflusses von Erinnerungsfaktoren sprechen.

Wir haben uns die in Tab. 17 dargestellten Mittelwerteveränderungen auf den fundamental messenden Skalen des AM-GITTERs nachträglich noch einmal genauer betrachtet und auf Signifikanz geprüft (SI/SPSS, T-TEST (PAIRS); vgl. BEUTEL & SCHUBÖ 1983, S. 261 f.). Dabei fällt auf, daß sich Veränderungen in den Mittelwerten, die

Signifikanz (zweiseitig) erreichen, durchgehend in einer Zunahme von Aggressionsmotiv- und in einer Abnahme von Aggressions-Hemmungs-Werten zwischen Erst- und Zweitmessung abbilden. So erhöht sich bei der Wiederholungsmessung nach 2 Wochen der Aggressionsmotivkennwert A von durchschnittlich 29.9 auf 27.2 Rohpunktwerte (sig. 2 %). Diese Zunahme geht offensichtlich auf die Anteile des Aggressionsmotivs zurück, die wir mit der DIFF-Skala ADH isolieren, auf der wir im gleichen Zeitintervall eine Zunahme von durchschnittlich 9.7 auf 13.3 Rohpunktwerte verzeichnen (sig. 2 %), wohingegen auf den Bedürfnis- und Erfolgserwartungs-bezogenen Motivanteilen (DIFF-Skala ADM) eine Veränderung von durchschnittlich 12.1 auf 14.4 Rohpunktwerte stattfindet (n.s.). Nach einem Zeitintervall von 4 Wochen vermindert sich der Kennwert für Aggressions-Hemmung (AH) von durchschnittlich 37.5 auf 29.8 Rohpunktwerte (sig. 2 %). Hier liegen die Verhältnisse jedoch wesentlich komplizierter. Auf der differentiellen Hemmungs-Skala AHI fällt der Mittelwert von 14.1 auf 11.2 (sig. 3 %), auf der differentiellen Hemmungs-Skala AHE von 15.2 auf 11.6 (sig. 2 %). Gleichzeitig steigt die Varianz ersterer von 50.0 auf 54.4, wohingegen die Varianz letzterer von 56.6 auf 50.8 fällt. Dabei erreicht nur die Veränderung in der Gesamtverteilung der Skala AHI Signifikanz (2 %; vgl. Tab. 18); es handelt sich hier offensichtlich um einen kombinierten Mittelwerte-Varianz-Effekt. Wir halten es für sinnvoll, auf diese Unterschiede ggfls. bei einer erneuten Durchführung der Stabilitätsanalyse zu achten, wollen sie jedoch vorerst nicht weiter interpretieren.

Wenn wir methodisch vergleichbar arbeitende Verfahren im Bereich der Persönlichkeitsdiagnostik von Kindern heranziehen, so liegen wir mit unseren Stabilitätskennwerten für das AM-GITTER praktisch in der erwarteten Größenordnung. PETERMANN & PETERMANN (1980) berichten für den EAS-J (vgl. Abschnitt 3.3 in Teil I dieser Arbeit) eine Wiederholungszuverlässigkeit von .71 nach 8 Wochen. Es handelt sich bei diesem Verfahren zwar nicht um einen Motivtest; es liegen jedoch Parallelitäten im thematischen Bereich (Aggression), in der konzeptuellen Gestaltung (abverlangt werden vom Probanden Reaktionen auf Wort-Bild-Kombinationen) und in der Population (Jungen zwischen 9 und 13 Jahren) vor. In ihrer Größenordnung entsprechen unsere Retestkoeffizienten auch etwa den von SCHMALT (1976 a, b) für das äquivalent arbeitende Leistungsmotiv-Gitter (LM-GITTER) für Kinder angegebenen. Hier liegen Parallelitäten in der Meßwerte-Ebene (Motive), in der konzeptuellen Gestaltung (GITTER) und in der Population (3. und 4. Klasse Grundschule) vor.

Übereinstimmend mit den Befunden von SCHMALT sowie PETERMANN & PETERMANN liegen unsere Stabilitäten über denjenigen, die für den TAT berichtet werden (vgl. SCHMALT 1973, S. 237 f.; 1976 a, S. 123 f.; McCLELLAND 1980, S. 34), sodaß wir insgesamt von einem wohl konstruktionsbedingten Stabilitätsgewinn gegenüber rein projektiv arbeitenden Verfahren ausgehen können.

3.3 Diskriminante und konvergente Validität

3.3.1 Vorüberlegungen

Die nachfolgende Untersuchung dient dem Zweck, erste Informationen über den Standort des AM-GITTERs einzuholen. Dazu sollen Korrelationsmuster mit validitätsnahen und -fernen Testvariablen aufgezeigt werden (konvergente bzw. diskriminante Validität; vgl. hierzu auch LIENERT 1969, S. 262; KOECK & STRUBE 1977, S. 194 ff.).

Formal rücken wir mit diesem Vorgehen in die Nähe des methodischen Ansatzes von CAMPBELL & FISKE (1959). Wir übernehmen deren (traditionell trait-theoretisches) Validierungskonzept jedoch nicht; die o.g. Schritte sollen, wie gesagt, mehr informatorische Funktion haben. Auch aus theoretischen Überlegungen heraus scheint uns die Logik der Beweisführung nach dem Ansatz von CAMPBELL & FISKE zur Überprüfung der Validität unserer Motivkennwerte weniger geeignet, denn

(a) aufgrund der streng theoriegeleiteten Konstruktion unseres AM-GITTERs sind allemal keine übergreifenden, sondern bestenfalls punktuelle Zusammenhänge mit Kennwerten aus anderen Aggressivitätsmeßverfahren zu erwarten.

(b) Neben der Theorieableitung ist weiter zu beachten, daß wir mit unserem Verfahren vorerst nur einen relativ engen Ausschnitt aus der Gesamtheit möglicher, feindseliger Motivzentrierungen operationalisiert haben, nämlich speziell den Bereich des überdauernden Motivsystems, der sich auf den sozialen Umgang mit Gleichaltrigen bezieht. Die Items/Skalen der (dann letztlich herangezogenen) Vergleichstests dagegen sind auf einen wesentlich breiter gefächerten (extensiveren) Hintergrund dieses Persönlichkeitsmerkmals hin angelegt.

(c) Gestützt wird die Annahme eines erschwerten Korrelationsnachweises zu konstruktnahen Testvariablen durch den empirischen Tatbestand, daß die überhaupt in Frage kommenden Vergleichstests ohnehin kaum korreliert sind, was darauf hinweist, daß ihnen eine gemeinsame, für uns als valides Bezugskriterium interpretierbare Dimension (Aggression) offensichtlich fehlt (vgl. ROHKAMM 1974; SELG 1974). Insofern ist ein Validitätsnachweis via Konvergenz schon aus formalmethodischen Gründen kaum wahrscheinlich.

(d) Ferner sind die psychometrischen Güteeigenschaften der Vergleichsmeßwerte selbst nicht ganz unproblematisch.

(e) Außerdem sind - und damit greifen wir Überlegungen von McCLELLAND (1951, 1980, 1981) auf - Unterschiede in den Meßwerte-Ebenen zu berücksichtigen. Vermutlich erfassen die in Betracht stehenden Verfahren mit ihren unterschiedlichen meßtechnischen Ansätzen voneinander unabhängige Typen von Dispositionsvariablen (Motive, Werte, Eigenschaften), was der unter Punkt (c) genannten empirischen Befundlage zwischen den einzelnen Verfahren durchaus gerecht werden würde.

Der Ansatz von McCLELLAND (Punkt e) ist unseres Erachtens deshalb von Bedeutung, weil er eine theoretisch differenziertere Formulierung von Zusammenhangs-Annahmen zwischen Aggressionsmeßwerten unterschiedlicher Erhebungsmodalitäten gestattet, als es bislang in der Validierungspraxis von Aggressionstests üblich ist (vgl. SELG 1974). Daneben sollte er unserer Meinung nach bei der Validierung von Konstrukt-Tests ganz allgemein Beachtung finden.

Nach McCLELLAND werden Motive i.d. Regel aus einem relativ spontanen

(operanten) Phantasieverhalten mittels TAT erfaßt. Werte, nach McCLELLAND sog. "Schemata" (d.h. internale, kognitive Repräsentationen von Wertvorstellungen), werden meist mittels inhaltlich fest vorgegebener (respondenter) Fragebögen ermittelt, und die Erfassung von Eigenschaften (traits) ist auf Beobachtung(en) habitueller Denk- und/oder Handlungsroutinen beschränkt.

McCLELLAND nimmt nun weiter an, daß die Meßwerte-Ebenen respondenter und operanter Testverfahren orthogonal aufeinanderstehen und ähnlich unterschiedlich gelagerten Persönlichkeitsdimensionen entsprechen. Respondente Verfahren arbeiten danach mehr auf der Ebene des Bewußten, operante mehr auf der Ebene des Unbewußten. Deshalb ist auch nicht zu erwarten (wie man das gemäß dem klassischen trait-theoretischen Ansatz tut), daß die unterschiedlichen Ebenen entsprechenden Dispositionsvariablen untereinander korrelieren.

Das läßt beispielsweise den Schluß zu, daß das Leistungsmotiv (n-ach.), das sich spontan in der Phantasie einer Person entfaltet (ein Motiv-Indikator), nicht äquivalent der Bedeutung ist, die dieselbe Person harter Arbeit beimißt (ein Indikator für ein Wert-Schema) und beide wiederum nicht äquivalent der Anstrengung sind, die sie typischerweise in Arbeitssituationen zeigt (ein Indikator für eine Eigenschaft). Konsistenzen (im Sinne der Korrelation mit einer gemeinsamen Persönlichkeitsdimension) sind nach McCLELLAND demzufolge nur zwischen Meßwerten aufgrund von Testmethoden des gleichen Antwortstils und auch dort noch nur unter

weiter eingeschränkten Bedingungen zu erwarten, d.h. nur an ganz bestimmten Stellen seines typologischen, 12 Kategorien umfassenden Systems der Persönlichkeitsmessung (vgl. McCLELLAND 1981, S. 95), im allgemeinen jedoch zwischen Meßwerten auf der Grundlage als unterschiedlich typisierter Methoden der Datenerfassung nicht. Diese Annahme wurde neuerdings explizit geprüft und konnte empirische Bestätigung finden (WILSON 1983; vgl. hierzu auch die Zusammenstellung der Befunde bei McCLELLAND 1980).

Zum Nachweis divergenter Validität des AM-GITTERs scheint die Ausgangslage zunächst weniger problematisch zu sein. Wir sehen es jedoch als sinnvoll an, auch hier theoriebezogen zu argumentieren und vorzugehen. Keinesfalls sollte man aufgezeigte Divergenzen allein schon aufgrund von Argumenten, wie sie dem oben dargestellten Ansatz zu entnehmen sind, allzuleicht entkräften können.

3.3.2 Vergleichstests

Die Suche nach geeigneten Aggressivitäts-Meßverfahren für Kinder im Alter von ca. 10 Jahren bleibt im Ergebnis enttäuschend (vgl. Abschnitt 3.3, Teil I dieser Arbeit). Dabei zeigt sich im einzelnen

- die sehr geringe Zahl an überhaupt zur Verfügung stehenden Testverfahren,
- die mangelnde theoretische Fundierung, empirische Verankerung und damit der unklare Gültigkeitsbereich der Verfahren,

- die unzureichende Berücksichtigung meß- und testtheoretischer Qualitätserfordernisse bzw. deren unklarer theoretischer Rückbezug,
- die mangelnde Berücksichtigung situationsspezifischer Bindungen des zu erfassenden Konstrukts.

Dieser unbefriedigende Zustand war ja ursprünglich mit ein Grund für die Entwicklung eines eigenen Verfahrens gewesen.

Nachdem somit die Ausgangsvoraussetzungen für eine Validierung durch Nachweis von Konvergenzeigenschaften des AM-GITTERs als auch für dessen Standortbestimmung denkbar ungünstig sind, entschließen wir uns nach eingehender Prüfung, folgende Vergleichstests für die Konvergenzprüfung auszuwählen:

1. PICTURE FRUSTRATION TEST (PFT) von ROSENZWEIG in der deutschsprachigen Adaptation von DUHM & HANSEN (1957),
2. PERSÖNLICHKEITSFRAGEBOGEN FÜR KINDER 9 - 14 (PFK) von SEITZ & RAUSCHE (1976), aggressionsrelevante Subskalen.

Beide Testverfahren wurden bereits in Teil I dieser Arbeit, Abschnitt 3.3 vorgestellt.

Zum Nachweis von Divergenzeigenschaften des AM-GITTERs wählen wir folgende Testverfahren:

1. LM-GITTER (SCHMALT 1976 b),
2. PERSÖNLICHKEITSFRAGEBOGEN FÜR KINDER 9 - 14 (PFK) von SEITZ & RAUSCHE (1976), nicht aggressionsrelevante Subskalen.

Für diese Auswahl sind im einzelnen folgende Gesichtspunkte maßgebend gewesen:

Nach den eingangs angestellten Überlegungen ist im Hinblick auf Konvergenzeigenschaften des AM-GITTERs, das testmethodisch eine Art Zwitterstellung zwischen operantem und respondentem Verfahrenstyp einnimmt, anzunehmen, daß positiv-korrelative Beziehungen allenfalls zu dem ähnlich typisierbaren PFT resultieren könnten. Daneben ist der PFT das bekannteste und verbreitetste Aggressivitäts-Meßverfahren für Kinder. Die Entsprechung zwischen den (phänomenologisch interpretierten) Inhalten der Meßelemente (Items) und der zugrundeliegenden Frustrationstheorie ist relativ eng. Die Auswertungsobjektivität ist befriedigend, und es liegen auch neuere Normen vor (RAUCHFLEISCH 1979). Der PFT repräsentiert neben dem Bereich "Interaktion mit Gleichaltrigen" den des sozialen Umgangs mit älteren Jugendlichen und Erwachsenen und läßt eine differentielle Auswertung zu.

Unter den deutschsprachigen Persönlichkeitsfragebögen scheinen uns die aggressionsrelevanten Subskalen des PFK noch am geeignetsten; daneben erlaubt der PFK die Erfassung weiterer Aspekte der kindlichen Persönlichkeit und bietet damit gleichzeitig Möglichkeiten zum Aufzeigen von Divergenzeigenschaften des AM-GITTERs.

Hinter der Auswahl des LM-GITTERs, das die Erfassung einer Hoffnungs- und zweier Furchtkomponenten der kindlichen Leistungsmotivation erlaubt und die testmethodische Vorlage für unser eigenes Verfahren darstellt, steht, kurz umrissen, folgende Überlegung: Aufgrund einer relativ engen (im Hinblick auf die übrigen Verfahren, die sich hier prinzipiell anbieten) formalen (typologischen

sensu McCLELLAND) Nachbarschaft der Meßwerte beider GITTER wollen wir uns den Nachweis divergenter Validität unseres GITTERs soweit als möglich erschweren, um ihn, falls er gelingt, als umso überzeugender betrachten zu können. Ausschlaggebend für Meßwertedivergenzen sollte - im theoretisch idealen Fall - nur ein Faktor sein, die Unterschiedlichkeit beider Motivsysteme.

3.3.3 Methode

Die Datenerhebung erfolgt im Juli 1982 an fünf saarländischen Grundschulen. Wir beziehen alle dort eingerichteten 4. Klassen, insgesamt 12, in die Untersuchung ein. Die Untersuchung dient primär der Erhebung von Außenkriterien (vgl. Abschnitt 3.4). Aus Zeitgründen können daher nicht alle Probanden mit allen Vergleichstests untersucht werden. Nähere Einzelheiten vgl. Abschnitt

Im einzelnen ergeben sich für die verschiedenen Vergleichstests folgende Beteiligungszahlen: PFT (n = 46), PFK (n = 38), LM-GITTER (n = 32).

Wir berechnen im PFT einen generellen Kennwert für Extrapunitivität (Kat.-E.), einen für Intropunitivität (Kat.-I.) und einen für Impunitivität (Kat.-M). Dazu benutzen wir die neuen Normdaten, die bei RAUCHFLEISCH (1979) angegeben sind (STANINE). Eine differentielle Auswertung, etwa nach den angegebenen Situationsbereichen (Interaktionen mit Gleichaltrigen bzw. Erwachsenen) ist vorerst nicht vorgesehen.

Den PFK werten wir lt. Manual aus und benutzen die Rohwerte.

Aus ökonomischen Gründen bringen wir nur eine Kurzfassung des LM-GITTERs zur Anwendung. Wir wählen dazu die Situationsbereiche MANUELL (Bildsituationen Nr. 1, 7 und 18) und SCHULE (Bildsituationen Nr. 3, 11 und 16). Wir werten lt. Manual aus und benutzen die Rohwerte.

Das AM-GITTER werten wir auf allen vorgeschlagenen Skalen aus und benutzen ebenfalls die Rohwerte.

Zum Nachweis konvergenter und divergenter Zusammenhänge in den Meßwertereihen berechnen wir SPEARMAN-Rangkorrelationen. Die Signifikanz der Koeffizienten prüfen wir zweiseitig mit dem t-Test von FISHER (CLAUSS & EBNER 1975, S. 274 f.). Das gewählte Programm führt jeweils eine Bindungskorrektur durch.

3.3.4 Präzisierung der Zusammenhangsannahmen

3.3.4.1 PFT-Variablen

Bei der von uns durchgeführten Untersuchung verteilen sich die Meßwerte für die einzelnen Richtungsdimensionen der Aggression im PFT wie folgt (vgl. Tab. 19); die Beziehungen zwischen den Kennwerten für die einzelnen Richtungsdimensionen sind in Tabelle 20 dargestellt.

Bei einem direkten Vergleich der im PFT und in unserem Verfahren operationalisierten Konstrukta kann man noch am ehesten mit einer

PFT-Variablen	$\bar{x}$	s	n
Kat.-E.	4.61	2.08	46
Kat.-I.	5.48	2.49	46
Kat.-M.	4.98	1.97	46

$\bar{x}$: arithm. Mittel

s : Standardabweichung

Tab. 19 Verteilungen der PFT-Variablen (SI/SPSS, CONDESCRIPTIVE)

rho	Kat.-E.	Kat.-I.	Kat.-M.
Kat.-E.		-.59 +	-.75 +
Kat.-I.			-.02
Kat.-M.			

+ : sig. 0.1 %

Tab. 20 Rangkorrelationen zwischen den PFT-Variablen (n = 46; SI/SPSS, NONPAR CORR)

positiven Korrelation zwischen dem PFT-Kennwert für Extrapunitivität (Kat.-E.) und unserem Aggressionsmotivkennwert A rechnen (12). Unser Hemmungskonstrukt AH hat dagegen schon vergleichsweise wenig mit einer nach innen gerichteten (und dazu noch frustrationsbedingten) Aggression zu tun - nach unserer Ansicht ist das eine konflikthafte, pathologische Erscheinungsform autoaggressiven Verhaltens - , und insofern ist ein Zusammenhang mit dem

PFT-Kennwert für Intropunitivität (Kat.-I.) kaum wahrscheinlich. Den PFT-Kennwert für Impunitivität (Kat.-M.) sollte man laut Manual (vgl. auch RAUCHFLEISCH 1979) als einen Indikator dafür verstehen, wie hoch die Bereitschaft einer Person bemessen ist, den frustrierenden Konflikt zu ertragen. Es gibt nun keinen vergleichbaren Kennwert im AM-GITTER; wir wollen ihn aber versuchsweise zu unserem Kennwert für Gesamt-Aggressivität (Ges.-A.) in Beziehung setzen, da dieser möglicherweise Aspekte einer überdauernden, konflikthaften Aggressionsmotivation beinhaltet. Kritisch ist jedoch anzumerken, daß für die Variable Kat.-M. keine Reliabilität nachgewiesen ist. Im Handbuch von RAUCHFLEISCH (1979) ist als Stabilitätskennwert eine Retestkorrelation von .27 (n.s.) angegeben; die interne Konsistenz dagegen beträgt .33 (sig. 0.1 %). Letztere ist jedoch - abgesehen von ihrer unzureichenden Größenordnung - , wie RAUCHFLEISCH (1980, S. 54) selbst eingesteht, bei einem Verfahren, bei dem man im Verlauf der Testdurchführung gerade Änderungen im Antwortverhalten erwartet, schon vom Ansatz her problematisch, da nicht durchgängig von zwei faktoriell gleichwertigen Testhälften ausgegangen werden kann. Hinsichtlich der DIFF-Skalen des AM-GITTERs formulieren wir keine Zusammenhangsannahmen.

Zum Schluß möchten wir darauf hinweisen, daß die unserer PFT-Korrelationsmatrix (vgl. Tab. 20) zugrundeliegende 2-dimensionale Struktur keine eindeutige Interpretation erlaubt, was im Hinblick auf das immer noch ungeklärte "Meßwerteebenen-Problem" im PFT äußerste Behutsamkeit und Zurückhaltung in der Deutung der Kennwerte angeraten erscheinen läßt (13).

3.3.4.2 PFK-Variablen

Die Meßwerte auf den aggressionsrelevanten PFK-Subskalen MD 1 (aggressives Bedürfnis nach Ich-Durchsetzung), MD 6 (Maskulinität der Einstellung) und VS 2 (fehlende Willenskontrolle) verteilen sich in unserer Untersuchung wie folgt (vgl. Tab. 21); die korrelativen Beziehungen zwischen diesen Meßwerten sind in Tabelle 22 dargestellt.

PFK-Variablen	$\bar{x}$	s	n
MD 1	8.66	4.52	38
MD 6	2.26	1.47	38
VS 2	10.74	5.21	38

$\bar{x}$: arithm. Mittel

s : Standardabweichung

Tab. 21 Verteilungen der PFK-Variablen (SI/SPSS, CONDESCRIPTIVE)

rho	MD 1	MD 6	VS 2
MD 1		.35 +	.60 +
MD 6			.32 +
VS 2			

+ : sig. 1 %

Tab. 22 Rangkorrelationen zwischen den PFK-Variablen (n = 38; SI/SPSS, NONPAR CORR)

Die in unseren Daten aufgefundene Korrelationsstruktur der PFK-Subskalen MO 1, MO 6 und VS 2 (vgl. Tab. 22) entspricht hinsichtlich Richtung und Signifikanz der Korrelationen den im Manual vorgegebenen Werten. Der Befund bestätigt die von den Testautoren vorgelegte Interpretation dieser Teilstruktur des PFK, der zufolge die drei o.g. Subskalen den gemeinsamen Faktor F-II-01 ("derb-draufgängerische Ich-Durchsetzung") laden. Zur Dimensionalität dieser Skalen läßt sich dennoch kaum etwas Verbindliches sagen, da die von den Testautoren vorgelegte Interpretation einerseits dazu keine bedeutsamen Anhaltspunkte liefert, andererseits keine empirischen Befunde angegeben sind, denen man Hinweise entnehmen könnte. Die 8 Item - Skala MO 6 ist zudem psychometrisch schwach (KR 20 = .66), was auch mit der sehr ungünstigen Rohwerteverteilung (vgl. Tab. 21) zusammenhängt, sodaß von daher etwas ungünstigere Voraussetzungen für das Aufzeigen von Meßwertekonvergenzen bestehen als bei den übrigen PFK-Subskalen. Zudem lassen einige der aggressionsnahen MO 6 - Items einen mehr attitüdentheoretischen Bezugsrahmen vermuten, was die Aussicht auf Konvergenz weiter verringert. Wir erwarten demzufolge noch am ehesten einen nachweisbaren Zusammenhang zwischen unserem Aggressionsmotivkennwert A und den Meßwerten der PFK-Subskala MO 1 - insbesondere auch im Hinblick auf die Itemzusammensetzung dieser Skala - , dagegen kaum Zusammenhänge zu den Subtest-Variablen MO 6 und VS 2. Hinsichtlich der DIFF-Skalen des AM-GITTERs (hier wären ADM und ADH relevant) formulieren wir ebenfalls keine Annahmen, mit auch auch aufgrund der vergleichsweise hohen wechselseitigen Korrelation.

3.3.4.3 LMG-Variablen

In der von uns untersuchten Stichprobe errechnen wir für den Vergleichstest LM-GITTER (LMG) die in Tabelle 23 dargestellten Kennwerte für die Meßwerteverteilungen; die Meßwerte der drei LMG-Subskalen HE 1 ("Konzept guter eigener Fähigkeit und erfolgszuversichtliche Bewältigung eher "schwieriger" Aufgaben"), FM 1 ("Konzept mangelnder eigener Fähigkeit und Initiation von Handlungen zur Abwendung von Mißerfolg") und FM 2 ("Furcht vor Mißerfolg"; vgl. SCHMALT 1976 b, S. 12) sind dabei wie der Tabelle 24 zu entnehmen interkorreliert.

LMG-Variablen	$\bar{x}$	s	n
HE 1	17.16	5.69	32
FM 1	13.59	6.31	32
FM 2	13.25	4.13	32

$\bar{x}$: arithm. Mittel

s : Standardabweichung

Tab. 23 Verteilungen der LMG-Variablen (SI/SPSS, CONDESCRIPTIVE)

rho	HE 1	FM 1	FM 2
HE 1		-.14	-.15
FM 1			.05
FM 2			

Tab. 24 Rangkorrelationen (alle n.s.) zwischen den LMG-Variablen (n = 32; SI/SPSS, NONPAR CORR)

Die in unseren Daten aufgefundene Korrelationsstruktur der LMG-Variablen (vgl. Tab. 24) entspricht genau der erklärten Absicht des Testautors, zwischen den Kennwerten HE 1, FM 1 und FM 2 insgesamt Nullkorrelation aufzuzeigen (SCHMALT 1976 b, S. 16 f.). Auch die zugehörigen Meßwerteverteilungen (vgl. Tab. 23) weisen auf gute diskriminative Eigenschaften dieser Skalen hin. Die Interpretation dieser Meßwerte kann also unproblematisch dem Manual entnommen werden. Wie bereits angedeutet, erwarten wir hier durchgehend Nullkorrelation zwischen den Meßwerten beider GITTER. Dennoch ist es nicht ausgeschlossen, daß aufgrund besonderer Bedingungen in der Genese beider Motivsysteme es (innerhalb einer bestimmten Gesellschaft/Kultur) zu einer gleichgerichteten Entwicklung und damit korrelierten Meßwerten kommen kann. Theoretisch ist die Beziehung zwischen Leistungsmotiv und Aggressionsmotiv bei Kindern noch weitgehend unklar. Vereinzelt findet man in der Literatur Hinweise auf eine negative Korrelation zwischen Aggression und Schulleistung (z.B. ERON et al. 1974) oder umgekehrt Hinweise auf eine positive Korrelation zwischen Aggression und Schullust (z.B. WAGNER & BAUMGÄRTEL 1978). Diese Befunde erlauben jedoch keinen Rückschluß auf korrelative Zusammenhänge zwischen den o.g. Motiven.

Zusammengefaßt ergibt sich damit folgende Ausgangslage: Nach dem traditionellen, trait-theoretischen Ansatz sensu CAMPBELL & FISKE hätte man zwischen allen Aggressionskennwerten der in die Untersuchung einbezogenen Verfahren positive Korrelationen zu erwarten. Dies kann u.E. jedoch sowohl aufgrund theoretischer Überlegungen als auch empirischer Befunde nicht als wahrscheinlich angesehen

werden. Dagegen sind in Anlehnung an die theoretischen Vorstellungen von McCLELLAND differenziertere Zusammenhangsannahmen formulierbar, die bei der typologischen Zwitter-Stellung unseres Verfahrens vorerst aber noch nicht als Hypothesen festgeschrieben werden sollen (14). Nicht, weil wir die Vorstellungen McCLELLANDs insgesamt schon für hinreichend erwiesen halten, sondern weil sie in ihrer Konsequenz den empirischen Gegebenheiten als auch unseren Vorstellungen einer theoriegeleiteten Validierung näherstehen, neigen wir, was unsere Erwartungen angeht, insgesamt den Überlegungen McCLELLANDs zu. Neben einem in der Validierungsphase von Konstrukttests üblichen Vorgehen entsprechen die hier vorgesehenen Maßnahmen nicht zuletzt einem mehr praktisch orientierten Informationsbedürfnis zu wissen, welche Beziehungen zwischen unserem Verfahren und im diagnostischen Alltag angewandten Aggressivitäts-Meßverfahren für Kinder existieren. Die vorliegende Untersuchung sollte somit insgesamt auch unter einem vorwiegend explorativen Aspekt betrachtet werden.

3.3.5 Ergebnisse

3.3.5.1 Vergleichsmessung PFT

In der Tabelle 25 sind die Korrelationen zwischen den Meßwerten des PFT und denen des AM-GITTERs aufgeführt. Die Korrelationen streuen alle um Null. Die einzige Ausnahme ($r_{M,ADH}$) wäre zwar, sollte sie sich im Wiederholungsfall bestätigen lassen, theoretisch sinnvoll, kann jedoch hier genauso gut auf Zufall beruhen.

Daneben ist dieser Zusammenhang nicht unter Kontrolle des Einflusses der beiden übrigen PFT-Konstrukta berechnet, sodaß wir hier von einer Interpretation absehen. Ansonsten können wir lediglich feststellen, daß die Korrelationen zwischen dem PFT-Kennwert für Extrapunitivität und den Meßwerten des AM-GITTERs tendenziell die erwartete Richtung anzeigen, ohne jedoch Signifikanz zu erreichen. Gleiches gilt für die Beziehungen zwischen PFT-Intropunitivität und unseren Hemmungs-Meßwerten.

rho	A	AH	n-A.	Ges.-A.	ADM	ADH	AHI	AHE
Kat.-E.	.22	-.16	.24	.05	.17	.25	-.17	-.12
Kat.-I.	.03	.18	-.13	.14	.01	.08	.12	.22
Kat.-M.	-.27	.04	-.18	-.16	-.18	-.35 +	.10	-.01

+ sig 2 %

Tab. 25 Rangkorrelationen zwischen den PFT-Variablen und den einzelnen Subskalen des AM-GITTERs (n = 46; SI/SPSS, NONPAR CORR)

3.3.5.2 Vergleichsmessung PFK

Der Tabelle 26 ist das Zusammenhangsmuster zwischen den einzelnen Subskalen des AM-GITTERs und den zum Konvergenznachweis ausgewählten PFK-Subskalen MO 1, MO 6 und VS 2 zu entnehmen. Die Korrelationen streuen alle ausnahmslos um Null; keine erreicht Signifikanz.

Wir verzichten hier auf die Darstellung der Zusammenhänge zwischen den Subskalen des AM-GITTERs und den zum Divergenznachweis ausgewählten restlichen Subskalen des PFK. Auch hier finden wir - bis auf ganz wenige, und deshalb nicht interpretationswürdige Ausnahmen - nur sehr geringe, nicht signifikante Korrelationen. Diese hatten wir allerdings erwartet.

rho	A	AH	n-A.	Ges.-A.	ADM	ADH	AHI	AHE
MO 1	.06	.03	.00	.06	.03	.15	.03	-.02
MO 6	-.08	.07	-.08	-.01	-.07	-.04	.06	.03
VS 2	.01	-.08	.05	-.05	.04	.06	-.05	-.06

Tab. 26 Rangkorrelationen (alle n.s.) zwischen den PFK-Variablen und den einzelnen Subskalen des AM-GITTERs (n = 38; SI/SPSS, NONPAR CORR)

3.3.5.3 Vergleichsmessung LMG

In der Tabelle 27 haben wir die Zusammenhänge zwischen den einzelnen Subskalen des AM-GITTERs und den zum Divergenznachweis gewählten Subskalen des LM-GITTERs zusammengestellt. Hier zeigen sich die erwarteten Nullkorrelationen bis auf die Ausnahme, daß unsere Kennwerte für Aggressions-Hemmung signifikant mit FM 1 (Konzept mangelnder eigener Fähigkeit) korreliert sind. Über diesen Zusammenhang kommt auch die Negativkorrelation zwischen Netto-Aggression (n-A.) und FM 1 zustande. Den differentiellen Skalen AHI und AHE ist zu

entnehmen, daß sowohl die intrinsischen als auch die extrinsischen Anteile der Aggressions-Hemmung an dieser Korrelation mit FM 1 teilhaben. Insgesamt sprechen die Ergebnisse jedoch dafür, daß wir mit unserem GITTER in einem eigenständigen Motivbereich messen.

rho	A	AH	n-A.	Ges.-A.	ADM	ADH	AHI	AHE
HE 1	-.11	-.20	.09	-.20	-.04	-.05	-.09	-.21
FM 1	.05	.44^{++}	-.41^{+}	.30	.01	.10	.40^{+}	.41^{+}
FM 2	-.10	.18	-.29	.04	.02	-.14	.06	.24

+ : sig. 2 %
++ : sig. 1 %

Tab. 27 Rangkorrelationen zwischen den LMG-Variablen und den einzelnen Subskalen des AM-GITTERs (n = 32; SI/SPSS, NONPAR CORR)

Zusammengefaßt zeichnet sich damit nach den ersten Zusammenhangsanalysen folgendes Bild ab: Zum Vergleichstest PFT, zu dem wir noch am ehesten Zusammenhänge erwartet hätten, können wir keine auffinden; ebensowenig zu den aggressionsrelevanten Skalen des PFK. Bei letzteren hatten wir sie jedoch sowieso kaum erwartet. Die zum Divergenznachweis ausgewählten Skalen haben - bis auf eine Ausnahme - die erwarteten Zusammenhänge erbracht. Aufgrund der der eingangs vorgetragenen Kritikpunkte als auch der aufgezeigten Meßwertezusammenhänge innerhalb der Vergleichstests kann man

allerdings auch dagegenhalten, daß wir mit unserem methodischen Ansatz zum Nachweis von Meßwertezusammenhängen sowohl unseren theoretischen Überlegungen als auch den empirischen Gegebenheiten in den Vergleichstests nicht ganz gerecht geworden sind, sodaß sich bei einer Verbesserung des methodischen Ansatzes zur Auswertung der erhobenen Information durchaus ein klareres (und vielleicht auch insgesamt günstigeres) Bild abzeichnen könnte. Diese Überlegungen veranlassen uns letztenendes dazu, die Überprüfung der Zusammenhänge zwischen den Meßwerten des AM-GITTERs und denen der ausgewählten Vergleichstests noch einmal - unter methodisch angemesseneren Voraussetzungen - in Angriff zu nehmen.

3.3.6 Weiterführende Analysen

3.3.6.1 Vorüberlegungen

Die bislang zum Nachweis von Konvergenz- und Divergenzeigenschaften des AM-GITTERs durchgeführten Analysen sind - unabhängig von den Ergebnissen - vergleichsweise restriktiv. Zunächst wird, auf Seite des zu prüfenden Verfahrens wie auf Seite des Vergleichstests, jeweils nur eine Testvariable in die Analyse einbezogen. Im Falle der Vergleichsvariablen kommt dann häufig noch hinzu, daß sie mit anderen kriterialen Variablen des Vergleichstests, aus dem sie entnommen ist, mehr oder minder stark korreliert ist, sodaß nicht die "reine", vom Konstrukt her intendierte Meßinformation verrechnet wird. Zudem werden in allen Fällen lineare Meßwertezusammenhänge vorausgesetzt.

Es ist evident, daß bei unserer problematischen Ausgangslage die o.g. Voraussetzungen umso stärker hinsichtlich ihres potentiellen Einflusses zu gewichten sind. Konkret bedeutet das, daß die Wahrscheinlichkeit von Meßwertekonvergenzen steigt, wenn wir unter weniger restriktiven Voraussetzungen analysieren. Dann ist aber auch gleichzeitig zu erwarten, daß die Wahrscheinlichkeit von Meßwertedivergenzen sinkt, was einer Erschwernis des Nachweises von Divergenzeigenschaften unseres Verfahrens gleichkommt. Letzteres entspricht jedoch gerade unserer Absicht.

Allgemein gesprochen wollen wir daher die Zusammenhangsanalysen vom univariaten auf den multivariaten und vom linearen auf den nichtlinearen Fall erweitern. Im Hinblick auf die gewählten Modellvorstellungen und die darauf beruhenden Analysetechniken sind die Zusammenhangsannahmen zwischen unserem Verfahren und den gewählten Vergleichstests dabei allerdings noch einmal neu zu überdenken.

Bei der Prüfung auf nichtlineare Zusammenhänge werden schon vom Modell her ungleich mehr Möglichkeiten überzufälliger Meßwertebeziehungen berücksichtigt. Insofern ist hier generell die Erwartung gerechtfertigt, daß - in Anlehnung an die eingangs bereits formulierten Zusammenhangsannahmen - die Voraussetzungen, Konvergenzen nachzuweisen, erleichtert, die Voraussetzungen dagegen, Divergenzen nachzuweisen, erschwert werden.

Bei der Kontrolle von Störvariablen auf Seite der Vergleichstests

kann man die gleichen Erwartungen formulieren. Es ist aufgrund der in den Vergleichstests bestehenden Korrelationsverhältnisse (vgl. Tab. 20, Tab. 22 und Tab. 24) jedoch nur sinnvoll, diesen Einflüssen im PFT als auch im PFK nachzugehen. Im Hinblick auf den Vergleichstest LM-GITTER wollen wir statt dessen prüfen, ob sich die zwischen FM 1 und unseren Meßwerten für Aggressions-Hemmung gefundene Beziehung auf eines unserer differentiellen Hemmungs-Konstrukta (intrinsische vs. extrinsische Hemmung) isolieren läßt.

Bei der Prüfung auf skalenübergreifende Zusammenhänge (gleichzeitige Berücksichtigung mehrerer Subtestvariablen auf Seite des AM-GITTERs wie auf Vergleichstest-Seite) verzichten wir von vornherein auf die zum Divergenznachweis ausgewählten PFK-Subskalen und konzentrieren uns statt dessen auf den Vergleichstest LM-GITTER. Beim PFT hatten wir sowieso nur punktuelle Zusammenhänge erwartet. Aufgrund dessen unklarer Dimensionalität kann eine Linearkombination mehrerer Richtungskennwerte für Aggression allenfalls sehr spekulativ interpretiert werden. Es ergeben sich von daher keine theoretischen Anhaltspunkte, von denen aus man einen übergreifenden Zusammenhang ableiten könnte. Wir erwarten demzufolge keinen übergreifenden Zusammenhang zwischen PFT und AM-GITTER. Beim PFK allerdings sind im Hinblick auf unser theoretisches Aggressionskonzept "günstige" Items ungleich über die Subskalen MO 1, MO 6 und VS 2 verteilt. Daher liegt die Vorstellung nahe, daß eine Linearkombination aus Meßwerten dieser drei Skalen die Aussicht auf Konvergenz zu unserem Testverfahren verbessert.

3.3.6.2 Methode

Zur Kontrolle des Einflusses der Konstrukta auf der Seite des Vergleichstests, auf die sich unser Zusammenhangsmaß im univariaten Fall nicht bezieht, berechnen wir multiple partielle Korrelationen. Deren Signifikanz prüfen wir zweiseitig nach dem t-Test von FISHER (CLAUSS & EBNER 1975, S. 274 f.).

Zur Überprüfung der Daten auf subskalenübergreifende Zusammenhänge zwischen AM-GITTER und Vergleichstest berechnen wir kanonische Korrelationen (vgl. GAENSSLEN & SCHUBÖ 1973; MOOSBRUGGER 1978). Das Modell der kanonischen Korrelation läuft hier auf die Überprüfung der Annahme hinaus, daß es im Raum gemeinsamer Faktoren des zum Vergleich anstehenden Testverfahrens einen speziell gerichteten Summenvektor (kanonische Variable) gibt, der mit einem entsprechenden Summenvektor der AM-GITTERs einen gemeinsamen Faktor lädt. Die kanonische Korrelation erfordert allerdings, ähnlich wie die Dimensionen im faktorenanalytischen Modell, eine theoretisch glaubwürdige Interpretation der kanonischen Variablen, falls damit ein Zusammenhang zwischen zwei Variablen-Sets (sinnvoll) belegt werden soll. Die Signifikanz prüfen wir hier über die chi^2-Approximation von WILK's Lambda (MOOSBRUGGER 1978, S. 119 ff.).

Der Frage nichtlinearer Zusammenhänge gehen wir mittels bivariater Streudiagramme ("Scatterdiagramme") nach.

Im übrigen beruhen die nachfolgenden Berechnungen auf den gleichen Meßwerten wie die vorangegangenen.

3.3.6.3 Ergebnisse

3.3.6.3.1 Vergleichsmessung PFT

PARTIELLE KORRELATIONEN: Wir kontrollieren nun auf PFT-Seite jeweils den Einfluß der beiden Aggressionsrichtungs-Konstrukta, zu denen wir gerade keine Konvergenz aufzeigen wollen, und wählen zudem ganz bestimmte Punkte aus, an denen wir Zusammenhänge zwischen den beiden Testverfahren erwarten/vermuten können. Das sind vor allem die Berührungsstellen Extrapunitivität /Aggressionsmotiv-Meßwerte sowie Intropunitivität/Meßwerte für die Aggressions-Hemmung, wobei wir hoffen, diese Zusammenhänge über die differentiellen Skalen des AM-GITTERs weitergehend aufhellen zu können. Das nach diesem Vorgehen berechnete Zusammenhangsmuster zwischen den beiden Testverfahren ist in Tabelle 28 dargestellt.

$r_{xy.z_1z_2}$	A	AH	n-A.	Ges.-A.	ADM	ADH	AHI	AHE
WITH E BY I,M	.42 +++		.31 +		.41 ++	.46 ++++		
WITH I BY E,M		-.04					-.05	.01
WITH M BY E,I				.18				

+ : sig. 4 % / ++ : sig. 0.6 %/ +++ : sig. 0.5 %/ ++++ : sig. 0.2 %
E : Kat.-E. / I : Kat.-I. / M : Kat.-M.

Tab. 28 Multiple partielle Korrelationen zwischen ausgewählten PFT-Variablen und Subskalen des AM-GITTERs (n = 46; SI/SPSS, PARTIAL CORR)

Wir finden nun die erwarteten Zusammenhänge zwischen der PFT-Richtungsdimension Extrapunitivität und unseren Meßwerten für das Aggressionsmotiv, wobei der Zusammenhang am deutlichsten mit der DIFF-Skala ADH nachgewiesen werden kann. Es gelingt jedoch nicht, den o.g. Zusammenhang auf die mehr motivzielbezogenen (ADM) oder die mehr handlungsbezogenen Anteile des Aggressionsmotivs (ADH) zu isolieren ($r_{ADM,\ Kat.-E.\ .I,M,ADH} = .03$, n.s.; $r_{ADH,\ Kat.-E.\ .I,M,ADM} = .22$, n.s.). Das zeigt, daß der bestehende Zusammenhang wohl weniger auf konstruktbedingte Gemeinsamkeiten zurückgeführt werden kann.

Unsere Meßwerte für Aggressions-Hemmung stehen dagegen auch hier, wie es sich schon im univariaten Fall gezeigt hat, mit der PFT-Richtungsdimension Intropunitivität in keinem systematisch-linearen Zusammenhang.

Nach der Logik der PFT-Signierung wird Impunitivität u.a. dann gescort, wenn keine eindeutige Zuordnung einer Antwort zu den beiden übrigen Richtungsdimensionen möglich ist, sodaß genau dann beide Aspekte in diesem Kennwert konfundiert werden. Das hätte zu einem positiven Zusammenhang zu unserem Kennwert für Gesamt-Aggressivität (Ges.-A.) mit beitragen können. Dieser zeigt sich jedoch hier nicht; statt dessen wird der im univariaten Fall eruierte Befund gestützt.

KANONISCHE KORRELATIONEN: Wir gehen nun der Frage nach, ob sich auf der Ebene kanonischer Variablen ein signifikanter linearer und interpretierbarer Zusammenhang zwischen den beiden Testverfahren aufzeigen läßt. Dazu geben wir in die Analyse die PFT-Variablen Kat.-E.,

Kat.-I. und Kat.-M. einerseits und die Variablen A, AH, ADM, ADH, AHI und AHE des AM-GITTERs andererseits ein. Das Ergebnis der kanonischen Analyse ist in Tabelle 29 dargestellt.

1. kanonische Korrelation	=	0.56
Eigenwert	=	0.31
WILK's Lambda	=	0.60
chi^2	=	20.16
df	=	18.00
Signifikanzniveau	=	0.32

Tab. 29 Ergebnis der kanonischen Analyse ausgewählter Variablen des PFT u. des AM-GITTERs (n = 46; SI/SPSS, CANCORR)

Wir finden also im Ergebnis keine signifikante kanonische Korrelation zwischen dem Vergleichstest PFT und unserem AM-GITTER.

SCATTERDIAGRAMME: Um die Zusammenhangsanalysen nicht dogmatisch auf den linearen Fall zu beschränken, erzeugen wir bivariate Streudiagramme über alle Kombinationen PFT (Kat.-E., Kat.-I., Kat.-M.) X AM-GITTER (A, AH, n-A., Ges.-A., ADM, ADH, AHI, AHE). Wir finden in diesen Diagrammen keine deutlich erkennbaren Hinweise auf nichtlineare Zusammenhänge zwischen den beiden Testverfahren (SI/SPSS, SCATTERGRAM).

Mit den Ergebnissen, so läßt sich zusammenfassend feststellen, werden unsere Erwartungen hinsichtlich des Auspartialisierens von PFT-Variablen im großen und ganzen bestätigt. Aufgrund der größeren, relativen Konstruktferne der PFT-Variable Kat.-I. kann

auch mittels der Berechnung partieller Korrelationen keine Konvergenz unserer Meßwerte für Aggressions-Hemmung nachgewiesen werden. Das gilt auch für die Berührungsstelle Gesamt-Aggressivität/Impunitivität zwischen den beiden Verfahren. Wie erwartet, besteht keine signifikante (und interpretierbare) kanonische Korrelation; ebensowenig finden wir nichtlineare Zusammenhänge zwischen PFT und AM-GITTER.

3.3.6.3.2 Vergleichsmessung PFK

PARTIELLE KORRELATIONEN: Wie schon beim PFT, so partialisieren wir nun auch bei den zum Konvergenznachweis ausgewählten PFK-Variablen MO 1, MO 6 und VS 2 jeweils die beiden aus, auf die sich der Zusammenhang nicht beziehen soll und berechnen danach die in Tabelle 30 dargestellte Matrix.

$r_{xy.z_1z_2}$	A	AH	n-A.	Ges.-A.	ADM	ADH	AHI	AHE
WITH MO 1 BY MO 6, VS 2	.02	.04	-.02	.04	-.10	.08	.06	.02
WITH MO 6 BY MO 1, VS 2	-.16	.06	-.16	-.05	-.19	-.13	.01	.09
WITH VS 2 BY MO 1, MO 6	.02	-.09	.09	-.06	.12	.01	-.10	-.08

Tab. 30 Multiple partielle Korrelationen (alle n.s.) zwischen ausgewählten PFK-Variablen und Subskalen des AM-GITTERs (n = 38; SI/SPSS, PARTIAL CORR)

Die partiellen Korrelationen streuen alle um Null, ohne daß eine

Signifikanz erreicht. Der Ansatz führt nicht zu einer Veränderung der im univariaten Fall ermittelten Befundlage.

KANONISCHE KORRELATIONEN: Wir setzen die Konvergenzprüfung auf der Ebene kanonischer Variablen fort und geben in die Analyse ein: auf PFK-Seite die Variablen MD 1, MD 6 und VS 2; auf Seite des AM-GITTERs die Subtestvariablen A, AH, ADM, ADH, AHI und AHE. Das Ergebnis der kanonischen Analyse ist in Tabelle 31 zusammengefaßt.

1. kanonische Korrelation	=	0.47
Eigenwert	=	0.22
WILK's Lambda	=	0.70
chi^2	=	11.43
df	=	18.00
Signifikanzniveau	=	0.88

Tab. 31 Ergebnis der kanonischen Analyse ausgewählter Variablen des PFK u. des AM-GITTERs (n = 38; SI/SPSS, CANCORR)

Wir finden im Ergebnis keine signifikante kanonische Korrelation zwischen dem Vergleichstest PFK und unserem Testverfahren.

SCATTERDIAGRAMME: Zuletzt prüfen wir auch hier auf nichtlineare Zusammenhänge und erzeugen zu diesem Zweck bivariate Streudiagramme über alle Kombinationen PFK (MD 1, MD 6, VS 2) X AM-GITTER (A, AH, n-A., Ges.-A., ADM, ADH, AHI, AHE). Wir finden in diesen Diagrammen keine deutlich erkennbaren Hinweise auf nichtlineare Zusammenhänge zwischen den beiden Testverfahren (SI/SPSS, SCATTERGRAM).

Zusammengefaßt bestätigen die Ergebnisse unserer weiterführenden Analysen insgesamt die schon vorliegenden Befunde, denen zufolge wir zwischen den aggressionsrelevanten Subskalen des PFK und unserem Verfahren keine Konvergenz aufzeigen können.

3.3.6.3.3 Vergleichsmessung LMG

PARTIELLE KORRELATIONEN: Wie schon erwähnt, bietet die Interkorrelationsstruktur der Subskalen des LM-GITTERs keine Handhabe für eine nachträgliche Auspartialisierung von LMG-Variablen. Wir gehen statt dessen hier der Frage nach, ob sich die aufgefundene Korrelation zwischen FM 1 und unserer Aggressions-Hemmung auf deren intrinsische oder extrinsische Anteile isolieren läßt. Ausgehend von der vorgestellten Interpretation des Konstrukts FM 1 (Konzept mangelnder eigener Fähigkeit) liegt die Vermutung einer Konvergenz mit den extrinsischen Bestandteilen der Aggressions-Hemmung nahe. Letztere beinhaltet u.a. Furcht vor Gegenaggression, die mit auch auf der Vorstellung mangelnder eigener Fähigkeiten, Aggression eigenständig zu initiieren und erfolgreich durchführen zu können, beruhen kann.

Die Meßwerte der DIFF-Skalen AHI und AHE sind in unserem Datensatz korreliert mit rho = .62 (sig. 0.1 %). Wir prüfen nun die Korrelation der Meßwerte beider Skalen unter wechselseitigem Auspartialisieren mit den Werten der LMG-Skala FM 1 und erhalten als Ergebnis: $r_{FM\ 1,\ AHI\ .\ AHE} = .20$ (n.s.) ; $r_{FM\ 1,\ AHE\ .\ AHI} = .23$ (n.s.)+. Beide Korrelationen sind danach nicht mehr signifikant. Wir können

---

+ : SI/SPSS, PARTIAL CORR

demzufolge die oben aufgeworfene Frage nur so beantworten, daß sich die aufgefundene Korrelation zwischen FM 1 und unserem Konstrukt Aggressions-Hemmung nicht auf dessen intrinsische oder extrinsische Anteile isolieren läßt. Wir sehen darin einen Hinweis darauf, daß wohl weniger inhaltliche Gemeinsamkeiten der Meßwerteinhalte hinter der aufgefundenen Korrelation stehen, was unserem Bemühen, Divergenz nachzuweisen, entgegenkommt.

KANONISCHE KORRELATIONEN: In die nun durchgeführte kanonische Analyse geben wir als Variablen-Sets ein: HE 1, FM 1, FM 2 einerseits und A, AH, ADM, ADH, AHI, AHE andererseits. Das Ergebnis der kanonischen Analyse ist in Tabelle 32 dargestellt.

1. kanonische Korrelation	=	0.56
Eigenwert	=	0.31
WILK's Lambda	=	0.54
chi^2	=	15.97
df	=	18.00
Signifikanzniveau	=	0.60

Tab. 32 Ergebnis der kanonischen Analyse ausgewählter Variablen des LM- und des AM-GITTERs (n = 32; SI/SPSS, CANCORR)

Wir finden keine signifikante kanonische Korrelation zwischen den Motivtests LM-GITTER und AM-GITTER. Dieses Ergebnis bestätigt unsere Erwartung hinsichtlich diskriminanter Validität des AM-GITTERs.

SCATTERDIAGRAMME: Zur Abrundung der nunmehr bestehenden Befundlage erzeugen wir bivariate Streudiagramme über alle Kombinationen LM-GITTER (HE 1, FM 1, FM 2) X AM-GITTER (A, AH, n-A., Ges.-A., ADM,

ADH, AHI, AHE). Wir finden in diesen Diagrammen keine deutlich erkennbaren Hinweise auf nichtlineare Zusammenhänge zwischen den Meßwerten beider Motivstests (SI/SPSS, SCATTERGRAM).

Zusammengefaßt bestätigen auch hier die Ergebnisse unserer weiterführenden Analysen unsere Erwartung, mit dem AM-GITTER in einem eigenständigen Motivbereich zu messen.

3.3.7 Diskussion

Die vorliegenden Ergebnisse entsprechen den Erwartungen, die wir mit der durchgeführten Untersuchung zum Aufzeigen konvergenter und divergenter Validität aus theoretischen Überlegungen heraus verbunden haben.

Hinsichtlich konvergenter Validität zeigen die Ergebnisse unserer Berechnungen, daß mit dem AM-GITTER ein eigenständiges Aggressivitäts-Meßverfahren vorliegt, das zu den herangezogenen Vergleichstests, mit denen ebenfalls diverse Aspekte kindlicher Aggressivität erfaßt werden sollen, kaum Beziehungen aufweist. Weder zum PFT noch zu den aggressionsrelevanten Subskalen des PFK lassen sich übergreifende Zusammenhänge nachweisen, die in Richtung einer gemeinsamen Persönlichkeitsdimension interpretierbar sind. Auch die Überprüfung auf einfache korrelative Zusammenhänge sowie auf nichtlineare Beziehungen ändert an dieser Befundlage nichts.

Unternimmt man jedoch den Versuch, Unterschiede in der theoretischen Konzeption der Vergleichstests als auch in der Stringenz ihrer

operationalen Qualität durch Auspartialisierung nicht in Betracht stehender Vergleichstestvariablen zu minimieren, so kommen punktuell schwache, aber dennoch signifikante und auch theoretisch sinnvolle Beziehungen zum Vorschein, wenn dazu noch die theoretische Nähe des Vergleichskonstrukts relativ groß und der Meßwertetyp (sensu McCLELLAND) vergleichbar ist. So zeigen multiple partielle Korrelationen zwischen unseren Kennwerten für das Aggressionsmotiv und der PFT-Variable Extrapunitivität erst bei Kontrolle der beiden übrigen PFT-Richtungsvariablen Intropunitivität und Impunitivität einheitlich signifikante Beziehungen an. Weitere Berechnungen machen allerdings die Annahme wahrscheinlich, daß dieser aufgefundene Zusammenhang weniger auf konstruktbedingte, inhaltliche Gemeinsamkeiten zurückzuführen ist, was unseren theoretischen Annahmen zum PFT entspricht und durch die Ergebnisse der kanonischen Analyse nachdrücklich bestätigt wird. Dagegen ergeben sich zum PFK bei ähnlichen Vergleichen in keinem Fall solche Beziehungen. Dafür kann man vielerlei Gründe in Betracht ziehen; es läßt sich jedoch nicht entscheiden, welche davon ausschlaggebend sind bzw. an welchen es insgesamt liegt. Während der PFT im wesentlichen auf zwei situative Bereichsklassen (wobei eine derjenigen unseres AM-GITTERs entspricht) zentriert und ein operantes Verfahren (was wir auch von unserem Verfahren annehmen) ist, sind die respondenten PFK-Subskalen auf einen sehr extensiven Bereichshintergrund bezogen (trait-Konzept) und auch weniger Aggressionstheorie-geleitet konstruiert (im Gegensatz zu unserem Verfahren als auch dem Frustrationstheorie-geleitet konstruierten PFT).

Mit um unsere Vorstellungen, unter welchen Bedingungen Konvergenz erwartet werden kann, weiter zu prüfen, könnte man die PFT-Variablen unter dem Aspekt einer situativen Annäherung an die Bereichsklasse des AM-GITTERs neu bestimmen, indem man einerseits die Informationsanteile verrechnet, die sich auf die PFT-Bereichsklasse "Kind-Kind-Interaktion" beziehen, und andererseits - zur Kontrolle - diejenigen, die in der PFT-Bereichsklasse "Kind-Erwachsenen-Interaktion" gesammelt werden. In den von uns verrechneten PFT-Variablen sind die beiden o.g. Bereichsklassen konfundiert, sodaß sie einen extensiveren Charakter aufweisen als unsere Meßwerte sowie die nach dem o.g. Vorschlag bereichszentriert bestimmten PFT-Werte. Obgleich damit formal noch nicht alle Möglichkeiten zum Aufzeigen konvergenter Beziehungen ausgeschöpft sind, halten wir ein weiteres Vorgehen in diese Richtung nicht für sinnvoll.

Diese Ergebnisse bestätigen nun unsere eingangs vorgetragene kritische Einstellung zu der Validierungsstrategie nach CAMPBELL & FISKE und stützen den hier in die Überlegungen mit einbezogenen Meßwerte-Ebenen-Ansatz sensu McCLELLAND, wonach Korrelationen zwischen verschiedenen Persönlichkeitstests, die auf unterschiedlichen Meßwerte-Ebenen arbeiten, kaum zu erwarten sind. Da nun das als Vergleichstest zum Nachweis divergenter Validität herangezogene LM-GITTER auf derselben Meßwerte-Ebene wie unser Verfahren arbeitet, dagegen in einem anderen Motivbereich mißt, spricht gerade für divergente Validität, wenn trotz gleicher Meßebene keine korrelative Beziehung auftritt, wofür unsere Befunde eigentlich recht deutlich sprechen.

Dennoch wurde eine Beziehung zwischen den beiden GITTERN an der LMG-Nahtstelle FM 1 festgestellt. Das Konstrukt FM 1 selbst weist allerdings eine Korrelation zum KAT (KINDER-ANGST-TEST von THURNER & TEWES 1969; vgl. BRICKENKAMP 1975, S. 447 f.) auf, und wir entnehmen einem Hinweis von SCHMALT (1976 a, S. 127), daß Skalen zur Erfassung von Angst erfahrungsgemäß und unabhängig von deren theoretischem Bezugssystem geringfügig positiv korreliert sind. Dafür spricht zunächst auch ein weiterer Detailbefund, demzufolge es uns nicht gelingt, die zwischen FM 1 und unseren Meßwerten für Aggressions-Hemmung bestehende Korrelation auf eines unserer beiden differentiellen Konstrukta für Aggressions-Hemmung zu isolieren. Jedoch ist diese Erklärung theoretisch nicht befriedigend. Daher sollten, zumal das Konstrukt "Angst" eine recht heterogene Größe darstellt (BIRBAUMER 1977) und verschiedene Tests zur Erfassung von Angst lediglich ganz bestimmte ihrer vielfältigen Aspekte erfassen (RAUCHFLEISCH 1980, S. 27 f.), zu weiteren Angst-Maßen Korrelationen errechnet werden, um den o.g. Befund ggfls. weiter aufhellen als auch die divergente Validität unserer Kennwerte für Aggressions-Hemmung umfassender abklären zu können. Entsprechende Vergleichsmessungen könnten evtl. mit Subskalen der Testverfahren KAT, AFS und HANES (vgl. BRICKENKAMP 1975) durchgeführt werden. Es ist allerdings auch denkbar, daß dennoch mit unseren Kennwerten für Aggressions-Hemmung - zumal in dem avisierten Bereich "Gleichaltrigenaggression" - auch die mangelnde Fähigkeit zur erfolgreichen Initiation und Durchführung von Aggression mit erfaßt wird.

Betrachten wir die aufgefundenen Zusammenhänge fundamental messender Skalen des AM-GITTERs zu den zum Konvergenznachweis ausgewählten Vergleichsskalen unter dem Aspekt der vektoriellen Zeigerstellung unserer Skalen im Generalfaktormodell, so können wir diese Überlegungen auf die Zusammenhänge unserer Kennwerte zu der PFT-Variablen Extrapunitivität (unter Kontrolle der beiden übrigen PFT-Variablen) beschränken.

Zunächst fällt auf, daß alle fundamentalen Meßwerte des AM-GITTERs, die den gleichen Ast der bipolaren Dimension SOZIALE WERTUNG laden, auch das gleiche Korrelationsverhalten gegenüber der PFT-Variablen Kat.-E. zeigen, und zwar dergestalt, daß sie mit dieser entweder signifikant positiv (Aggressionsmotiv-Meßwerte) oder nullkorreliert sind (Meßwerte für die Aggressions-Hemmung).

Da nun unsere Meßwerte für Aggressions-Hemmung keine Beziehung zur PFT-Variablen Kat.-E. aufweisen, können wir diesen Detailbefund mit auch als einen Hinweis auf divergente Validität selbiger auffassen, da sie als von den Aggressionsmotiv-Variablen unabhängige Variablen konzipiert sind und letztere ja durchgehend Korrelationen im Sinne konvergenter Validität zu Kat.-E. aufweisen.

Sodann entspricht es unserem (gegenwärtigen) theoretischen Verständnis der DIFF-Skalen ADM und ADH, daß letztere (tendenziell) etwas höher mit Kat.-E. korreliert als erstere, wenngleich auch die Unterschiede sehr schwach ausgeprägt sind und damit nur andeutungsweise bestehen. Das heißt nun im einzelnen, daß die Skala, in der wir die

auf ein Bedürfnis zur feindseligen Aggression unmittelbar verweisenden Bestandteile unseres Aggressionsmotiv-Konstrukts zusammengezogen haben, weniger mit Kat.-E. korreliert als diejenige, mit der wir die Aggressionshandlung selbst unmittelbar, ohne genauere Anzeichen bezüglich ihres motivationalen Hintergrunds darstellen. Bezüglich der vektoriellen Zeigerstellung der beiden DIFF-Skalen ADM und ADH deutet sich damit an, daß mit zunehmender Neigung eines Subtestvektors hin zur thematischen Dimension AGGRESSION die Korrelation zu Kat.-E. steigt, was genau dann einen Sinn ergibt, wenn in die Ausprägung einer Person auf der Dimension Kat.-E. Aspekte ihres (offen) aggressiven Verhaltens eingehen und weniger die ihrer überdauernden, feindseligen Motivation.

Ob letzteres zutrifft, läßt sich jedoch nicht ohne weiteres sagen, denn bis heute ist unklar, auf welcher Ebene der PFT mißt (vgl. SCHMIDTCHEN 1975, S. 218). ROSENZWEIG selbst läßt unterschiedliche Ebenen in der Interpretation der Meßwerte des PFT offen, RAUCHFLEISCH (1979, 1, S. 170, S. 188 f.) hält es auf empirischer Grundlage für entscheidbar. Er glaubt, daß mit Hilfe einer konfigurationsfrequenzanalytischen Behandlung der PFT-Daten aus der jeweiligen Kombination der über- und unterrepräsentierten PFT-Variablen die angesprochene Ebene gesichtet werden kann (ebd., S. 170). Interessant ist bei der von ihm in diesem Zusammenhang angeführten Untersuchung für uns hier nicht so sehr, daß danach eine Meinungs- und eine Verhaltens-Ebene als Interpretationsalternativen des Testverhaltens zur Wahl stehen, sondern daß die motivationstheoretisch

relevantere Ebene, die des "implicit behavior" (ebd., S. 76 f.), KFA-technisch nicht nachgewiesen werden kann (ebd., S. 170). Letztere entspricht der Ebene, in der projektive Verfahren messen (ebd., S. 77), und man könnte somit durchaus zu der Annahme gelangen, daß der PFT gerade in der Ebene, in der Motive (erfolgreich) gemessen werden, ein ungeeignetes Verfahren ist. Diese Annahme würde sich auch mit Befunden decken, denen zufolge bei einer gemeinsamen Faktorisierung von PFT- und TAT-Daten der Aggressivität zwei Methodenfaktoren, ein PFT- und ein TAT-Faktor (vgl. die Untersuchung von SELG & LISCHKE 1966; zit. n. SELG 1974), in Erscheinung treten.

Insofern wäre also, wenn wir mit der Skala ADM gegenüber der Skala ADH mehr die eigentlichen, intrinsischen Anteile des Aggressionsmotivs erfassen und die oben aufgeführte Argumentation zusätzlich zutrifft, dieser minimale Korrelationsunterschied plausibel.

Die zuletzt angestellten Überlegungen sind allerdings rein spekulativ und sollten u.a. noch einmal die theoretische Problematik verdeutlichen, mit der die hier durchgeführten Validierungsschritte belastet sind.

3.4 Korrelation mit Außenkriterien

3.4.1 Vorüberlegungen

Nachdem jetzt erste Informationen kriterienbezogener Validität unseres Testverfahrens vorliegen - wir haben dazu als Kriterien die Meßwerte von Vergleichstests gewählt -, wollen wir die Validierungsmaßnahmen nun dergestalt fortsetzen, daß wir als nächstes die Beziehungen unserer Meßwerte zu sog. "Außenkriterien" untersuchen, von denen wir annehmen, daß sie Aspekte offen aggressiven und aggressionsgehemmten Verhaltens vergegenständlichen.

Ein Hauptproblem bei diesem Vorgehen besteht auch hier zunächst darin, geeignete, d.h. reliable und valide Außenkriterien überhaupt erst zu finden (vgl. z.B. LIENERT 1969, S. 255 ff.; RAUCHFLEISCH 1980, S. 57). THORNDIKE (1949; zit. n. LIENERT 1969, S. 256 f.) spricht in diesem Zusammenhang von vorläufigen, Zwischen- und endgültigen Kriterien, die in dieser Abfolge eine sukzessive Approximation an den ins Auge gefaßten Sachverhalt darstellen und, je nachdem, die aufgefundenen Zusammenhänge entsprechend relativieren.

Wir wollen zur Erhebung von Außenkriterien ein Gleichaltrigen-Rating zum Einsatz bringen. Das scheint angezeigt, weil die situative Bereichsklasse des AM-GITTERs auf motivspezifische Aspekte feindseligen Erlebens, Deutens und sonstigen Verhaltens, das Jungen bei verschiedenen sozialen Anlässen im Umgang mit Gleichaltrigen zeigen, abgestimmt ist. Sodann lassen die positiven Ergebnisse,

die OLWEUS (1969, 1978) mit einem solchen Erhebungsansatz berichtet, erwarten, daß mittels so erhobener Kriterien valide Information zu erbringen ist. Die Grundkonzeption von OLWEUS soll deshalb für die Gestaltung unseres eigenen Ratings Verwendung finden. Dennoch kann man bei Kindern, die als Rater gänzlich untrainiert sind (eine spezielle Rater-Schulung ist dabei nicht vorgesehen), nicht ausschließen, daß sie in ihren Einschätzurteilen zu einer Reihe systematischer Fehler (vgl. GUILFORD 1954, S. 278 ff.) neigen, sodaß die darauf beruhenden kriterialen Verhaltensvariablen für die eingeschätzten Personen - mit auch unter Berücksichtigung der Güte des vorliegenden Instruments zur Einschätzung - nur ganz vorläufiger Art sind. Der Idealfall, daß ein Rating-Kennwert nur Information über den Eingeschätzten enthält, ist also zunächst unwahrscheinlich; man muß vielmehr damit rechnen, daß er auch Information beinhaltet, die auf Personcharakteristika des Raters verweist.

Solange also bei einem Rating-Verfahren die Annahme, daß die Einschätzwerte, die die Personen auf den einzelnen Skalen erhalten, in Abhängigkeit von verschiedenen Persönlichkeitsmerkmalen der Einschätzer systematisch verfälscht werden, nicht gänzlich ausgeschlossen werden kann, liegt in der Vorstellung, daß sich die Rating-Daten grundsätzlich auf zwei unterschiedlichen Ebenen interpretieren und damit auch analysieren lassen, eine gewissermaßen dritte, integrative Möglichkeit zur Verwertung der erhaltenen Daten. Wenn man nun die Aggressivität des Raters, was naheliegend ist, beim Einschätzprozeß in Rechnung stellt, ist eine darauf bezogene Überprüfung der

Außenkriterien jedoch schon ein unmittelbar konstruktbezogener Validierungsschritt, der in der Darstellung einen eigenen Abschnitt erfordert.

Üblicherweise erwartet man nun zwischen den Meßwerten eines zu prüfenden Meßverfahrens und unabhängig davon erhobenen Außenkriterien (meist positive) signifikante, korrelative Zusammenhänge, was formal auf einen Konvergenznachweis hinausläuft. Damit ist im Grunde die gleiche Situation gegeben wie bereits bei dem Versuch, konvergentes Verhalten der Meßwerte des AM-GITTERs zu ganz bestimmten, ausgewählten Vergleichsmaßen nachzuweisen. Wir haben dies im Abschnitt 3.3 getan und dort in unsere Überlegungen die Theorie McCLELLANDs einbezogen, die genauere Spezifikationen zu den Konvergenzbedingungen/-voraussetzungen erlaubt. Es spricht nichts dagegen, sie auch hier zur Anwendung zu bringen. Unter den geschilderten Voraussetzungen, unter denen wir in die Analyse von Rating-Daten einsteigen, ist ihre Anwendung vielmehr sinnvoll, denn nach McCLELLAND sind bei einer Analyse der Daten

(a) auf Ebene der Einschätzer (RATER-Ebene) TOT (thought operant trait) - Daten,

(b) auf Ebene der Eingeschätzten (RATEE-Ebene) dagegen TRT (thought respondent trait) - Daten zu erwarten.

In Anwendung der o.g. Theorie und unter den genannten Voraussetzungen, daß die Rating-Daten sowohl Information über den Eingeschätzten als auch über den Einschätzer enthalten, sind nun die Voraussetzungen zum Konvergenznachweis unter einer weiteren Voraussetzung (Datentyp der Meßwerte des AM-GITTERs) auf den beiden sich

anbietenden Ebenen der Analyse unterschiedlich günstig und damit Zusammenhänge unterschiedlich wahrscheinlich. Unter der Annahme, daß unser Verfahren mehr operante, einem TAT vergleichbare Meßwerte liefert, so sind bei einer Analyse der Daten auf RATEE-Ebene (wie vorgesehen) kaum Beziehungen zu erwarten; nimmt man die systematische Verfälschung im Einschätzprozeß selbst dagegen als kriteriale Variable (RATER-Ebene), so wird aufgrund typologischer Äquivalenz der Meßwertereihen ein Zusammenhang wahrscheinlich. Unter der Voraussetzung jedoch, daß die Meßdaten des AM-GITTERs mehr respondenten Charakter besitzen, sind in genau umgekehrter Richtung Zusammenhänge zu erwarten, d.h. die Bedingungen für eine Meßwertekonvergenz zwischen den Werten einer Person im AM-GITTER und ihren in den Außenkriterien erhaltenen günstig.

Die letztgenannten Überlegungen bieten u.E. Anhaltspunkte dafür, die hier vorgesehenen Schritte - je nach Ergebnislage - sinnvoll weiter fortsetzen zu können, weswegen wir sie an dieser Stelle auch vorgetragen haben.

3.4.2 Ratingmethode

3.4.2.1 Entwicklung

Wir wollen mittels eines Gleichaltrigenratings in erster Linie Aspekte aggressiven und aggressionsgehemmten Verhaltens von Kindern erfassen, um für die Aggressionsmotiv- als auch für die Aggressions-Hemmungs-Meßwerte des AM-GITTERs jeweils ein unabhängiges

Außenkriterium zu erhalten. Bei der Heranziehung von Gleichaltrigen als Rater können wir davon ausgehen, daß sie die einzuschätzenden Personen (es ist geplant, die Untersuchung in Schulklassen durchzuführen) gut kennen, gerade auch in Bezug auf diejenigen Aspekte des Verhaltens, wie wir erfassen wollen. Um den Charakter dessen, was als eigentliche Dimension (feindselige Aggression) interessiert, nicht zu dominant werden zu lassen, lassen wir dazu noch Einschätzungen auf einer Puffer-Dimension (prosoziales Verhalten) vornehmen. Da wir zusätzlich in den Außenkriterien einige Aspekte aggressiven und aggressionsgehemmten Verhaltens differenzieren, umfaßt unser System zur Verhaltenseinschätzung im ersten Entwurf sieben verschiedene Rating-Skalen, die sich auf die drei o.g. Verhaltensdimensionen verteilen.

Formal sehen wir für jede Skala ein gesondertes Blatt (30 cm x 42 cm, Querformat) vor. Oben auf dem Blatt steht die Fragestellung (Statement), darunter befinden sich zwei Reihen von je sieben gleichgroßen Quadraten. Die erste dient zur Darstellung der Metrik (verbal und mit Zahlen), die zweite zur Einordnung der Ratees (vgl. ANHANG V).

Die Skalen werden in jeweils gleichbleibender Reihenfolge (ihrer Numerierung entsprechend durchlaufend von 1 bis 7; vgl. Tab. 33) den Ratern im Einzelversuch vorgelegt, dazu die Namen der Ratees, die auf gut durchmischten Kärtchen stehen. Der Rater selbst ist dabei stets ausgeschlossen (kein Selbstrating), um beim Einschätzprozeß nicht auf seine Person als Bezugsmaßstab hingewiesen zu werden. Durch diese Vorkehrung erhoffen wir uns eine höhere Objektivität der Daten.

Die Bezugsnorm für die quantitative Ausprägung des jeweils einzuschätzenden Verhaltensaspekts soll die Gruppe liefern, die vom Rater eingeschätzt wird und der er selbst angehört.

Die Aufgabe des Raters besteht im einzelnen darin, die Kärtchen mit den Namen in die leeren Kästchen unter den zutreffenden Ausprägungsgrad für das in Frage stehende Verhaltensmerkmal zu verteilen. Dabei hat er freie Wahl, wieviele Kästchen er besetzen und auch, wieviele Kärtchen er in ein Kästchen legen will. Darüber hinaus wird ihm mitgeteilt, daß er die relative Position schon eingeschätzter Mitschüler während der Aufgabe noch ändern kann. Nach der Einordnung aller Ratees wird er aufgefordert, seine Verteilung noch einmal zu überprüfen und ggfls. noch einmal zu verändern. Bei den etwas längeren Statements wird durch zwanglose Rückfrage sichergestellt, ob er diese auch sinngemäß verstanden hat.

Zur Verbesserung der Reliabilität der Rating-Daten wählen wir per Zufall aus jeder Klasse vier Jungen als Rater aus und versichern uns zuvor der Bereitschaft, am Versuch gerne teilnehmen zu wollen. Die Werte für die Außenkriterien selbst berechnen wir nach dem Mittelungs-Schätzverfahren (vgl. LIENERT 1969, S. 276). Das impliziert den Vorteil, über eine kleine Rater-Stichprobe zu verfügen und damit die Daten auf RATER-Ebene untersuchen zu können. Demzufolge erhält also jeder Ratee als Kennwert für ein Außenkriterium das arithmetische Mittel vier unabhängiger Einzelschätzungen, die jeweils zwischen 1 und 7 variieren können.

3.4.2.2 Skalen

In der Tabelle 33 haben wir die Skalen, mit denen wir die Außenkriterien erheben, zusammengestellt.

RATING-SKALA (Statement)	VORGESEHENE THEMATISCHE DIMENSION	BESONDERE ASPEKTE
(1) Er teilt mit anderen, wenn er etwas hat.	Prosoziales Verhalten	Puffer
(2) Tut absichtlich nicht, was der Lehrer sagt.	Aggression	feindselige Aggression; Situationsbereich: Interaktion mit Autoritätspersonen;
(3) Traut sich nicht, sich zu wehren, auch wenn er sich ärgert.	Aggressions-Hemmung	angstgetönte Aggressions-Hemmung (extrinsische AH)
(4) Er hilft anderen.	Prosoziales Verhalten	Puffer
(5) Fängt mit anderen Raufereien an.	Aggression	feindselige Aggression; Situationsbereich: Interaktion mit Gleichaltrigen;
(6) Hat ein schlechtes Gewissen, wenn er frech war.	Aggressions-Hemmung	schuld- und schamgetönte Aggressions-Hemmung (intrinsische AH)
(7) Wehrt sich, wenn er angegriffen wird.		instrumentelle Aggression i. Sinne einer nicht aggressionsmotivspezifischen, habituellen Wehrbereitschaft;

Tab. 33 Skalen zur Erhebung der Gleichaltrigenratings (Außenkriterien)

Wir wollen dabei speziell mit der Skala Nr. 2 die auf eine situative Bereichsklasse bezogenen Verhaltenskorrelate feindseliger Aggression erfassen, die im AM-GITTER nicht berücksichtigt

ist. Eventuell lassen sich Meßwertezusammenhänge mit diesem Außenkriterium auffinden, die erste Annahmen zur Generalisierbarkeit bzw. zum Extensitätscharakter der Meßwerte des AM-GITTERs erlauben. Mit Skala Nr. 7 soll Information über eine habituelle Wehrbereitschaft erhoben werden, die p.d. mit Aggression in unserem Verständnis nichts zu tun hat. Zu diesem Außenkriterium sollen unsere Aggressionsmotiv-Meßwerte weniger eine Beziehung aufweisen, dagegen zu dem mit der Skala Nr. 5 erhobenen. Mit den Skalen Nr. 3 vs. Nr. 6 soll ein Differenzierungsaspekt innerhalb der Aggressions-Hemmung nun auch auf Ebene der Außenkriterien eingebracht werden, der in der Konstruktion des AM-GITTERs berücksichtigt wurde und in der Folge zur Entwicklung zweier differentieller Skalen für das überdauernde Motiv Aggressions-Hemmung führte.

Bei der von uns durchgeführten Untersuchung verteilen sich die Meßwerte auf den einzelnen Rating-Skalen wie in Tabelle 34 dargestellt.

RATING-Variablen	$\bar{x}$	s	n
(1) TEILEN	4.11	1.24	77
(2) AUTAGG	3.60	1.41	77
(3) TRAUTN	3.13	1.38	77
(4) HELFEN	4.19	1.29	77
(5) RAUFEN	3.57	1.48	77
(6) GEWISS	3.61	1.19	77
(7) WEHREN	5.03	1.39	77

$\bar{x}$: arithm. Mittel; s : Standardabweichung

Tab. 34 Verteilungen der Rating-Variablen (SI/SPSS, FACTOR)

3.4.2.3 Dimensionalität der Skalen

Da es sich bei der von uns entworfenen Ratingmethode zunächst um ein noch gänzlich ungeprüftes Instrument handelt, wollen wir, bevor wir konkret in die Zusammenhangsanalysen einsteigen, fürs erste die interne Struktur der Daten, die es liefert, prüfen, und dabei feststellen, ob sie den vorgesehenen thematischen Dimensionen, auf denen wir Außenkriterien erheben möchten (vgl. Tab. 33), entspricht.

Dazu untersuchen wir die mit unseren Skalen erhobenen Kennwerte für die Ratees im faktorenanalytischen Modell (THURSTONE 1969). Im Ergebnis akzeptieren wir die varimaxrotierte Dreifaktorenlösung (9 Iterationen), die die Zusammenhänge zwischen den Ratings am überzeugendsten beschreibt (vgl. Tab. 35).

Wir umschreiben die mit dem Faktor 1 zum Ausdruck gebrachte bipolare Dimension kurz mit dem Begriffspaar "WEHRBEREITSCHAFT vs. WEHRANGST". Bei Faktor 2 handelt es sich um den unipolaren Faktor "PROSOZIALES VERHALTEN" und bei dem unipolaren Faktor 3 um eine Dimension, die wir "FEINDSELIGE AGGRESSION" nennen wollen.

Um die Zuverlässigkeit unserer Dreifaktorenlösung in erster Näherung zu prüfen, wagen wir, trotz des relativ geringen Stichprobenumfangs, eine Zufallsteilung der Stichprobe in die zwei Teilstichproben S1 (n = 39) und S2 (n = 38) und rechnen sie dort jeweils noch einmal durch. Wir finden sie dabei in S1 nach 4 Iterationen voll bestätigt. Die Faktorenkonstellation ist die gleiche, die

ROTIERTE FAKTOREN

VARIABLE	Faktor 1	Faktor 2	Faktor 3	h^2
TEILEN	-.09	.80	-.03	.65
AUTAGG	.09	-.14	.62	.41
TRAUTN	-.78	-.03	-.07	.61
HELFEN	.07	.79	-.21	.67
RAUFEN	.48	-.06	.79	.86
GEWISS	-.29	.13	-.15	.12
WEHREN	.95	.09	.28	.98
Summe a_j^2	1.85	1.31	1.16	
% gemeinsame Varianz :	42.82	30.32	26.85	
% Gesamt-Varianz :	26.43	18.71	16.57	SUMME: 61.71

Tab. 35 Ladungsmatrix der Rating-Skalen (n = 77; SI/SPSS, FACTOR (PA 2))

Variable GEWISS hat eine Ladung von -.47 auf der Dimension WEHRBEREITSCHAFT vs. WEHRANGST (im Vergleich dazu: TRAUTN = -.56). Nach 8 Iterationen finden wir die varimaxrotierte Dreifaktorenlösung in S2 bestätigt. Es liegt wiederum die gleiche Faktorenkonstellation vor. Die Variable GEWISS weist hier allerdings eine Ladung von -.49 auf der Dimension FEINDSELIGE AGGRESSION auf, die man demzufolge hier auch als bipolar auffassen könnte (im Vergleich: TRAUTN = -.79 auf der Dimension WEHRBEREITSCHAFT vs. WEHRANGST).

Bedingt durch den Umstand, daß die Variable TRAUTN den Gegenpol zu der Variablen WEHREN bildet, ergibt sich kein deutlich in Erscheinung tretender Anknüpfungspunkt für eine potentielle Kovariation mit der Variablen GEWISS, die demzufolge innerhalb dieses eng umgrenzten Ratingverfahrens nicht mehr dimensional aufgeklärt werden kann. Darauf beruht auch ihre unzureichende Kommunalität. Auf der anderen Seite entspricht die mangelnde Kovariation zwischen den Variablen TRAUTN und GEWISS Dimensionierungsbefunden im AM-GITTER, wo wir auf Statement-Ebene zwei einander orthogonal zugeordnete Dimensionen für Aggressions-Hemmung (intrinsische und extrinsische Hemmung von Aggression) gefunden haben, die offensichtlich stabil sind. Daher kann man die mit der Ratingskala TRAUTN erfaßte Wehrangst vorerst durchaus im Sinne unserer extrinsischen Aggressions-Hemmung interpretieren. Die Befundlage spricht auch nicht dagegen, die mit der Variablen GEWISS erfaßten Verhaltensmerkmale als Hinweis auf ein Persönlichkeitsmerkmal zu verstehen, das mit unserem Konstrukt "intrinsische Aggressions-Hemmung" gleichgesetzt werden kann (ein "schwacher" Hinweis darauf ist die leicht negative Korrelation mit der Dimension FEINDSELIGE AGGRESSION in S2). Unter günstigeren Voraussetzungen in der Variablenbesetzung des Ratingverfahrens würde sich die Dimensionalität der Variablen GEWISS aller Wahrscheinlichkeit nach näher bestimmen lassen.

Die Variable RAUFEN erweist sich als faktoriell komplex und scheint mit auch Anteile einer habitualisierten Wehrbereitschaft zu umfassen.

Demgegenüber ist die Variable AUTAGG einfachstrukturiert und eindeutig durch die Dimension FEINDSELIGE AGGRESSION bestimmt. Aufgrund ihrer faktoriellen Komplexität halten wir die Variable RAUFEN als Außenkriterium für weniger geeignet; für ihre Verwendung als Außenkriterium spricht ihre hohe Ladung auf der Dimension FEINDSELIGE AGGRESSION sowie ihre Kongruenz mit dem Situationsbereich des AM-GITTERs (15).

Kurz zusammengefaßt läßt sich unter dimensionsanalytischen Gesichtspunkten folgendes resümieren: Mit Ausnahme der Skala GEWISS (niedrige Kommunalität) sind alle übrigen Ratingskalen innerhalb des vorgestellten dreidimensionalen Modells interpretierbar und - mit Ausnahme der Skala RAUFEN - auch eindeutig durch jeweils eine Dimension bestimmt. Das Modell scheint stabil und läßt den Schluß zu, daß die Rater auf den einzelnen Dimensionen unabhängig und damit strukturell valide geurteilt haben. Der Befund erhöht die Vertrauenswürdigkeit der Ratings als Außenkriterien.

3.4.3 Zusammenhangsannahmen

Nach Ermittlung der interdependenten Beziehungen zwischen den einzelnen Außenkriterien und deren dimensionsanalytischer Prüfung können wir unsere während der Entwicklung der Ratingmethode formulierten Vorstellungen bezüglich aufzuzeigender Zusammenhänge im wesentlichen beibehalten. Es ist nun gleichermaßen eine Korrelation zwischen unseren Meßwerten für das Aggressionsmotiv (A, ADM, ADH, n-A.)

und den Außenkriterien AUTAGG sowie RAUFEN zu erwarten. Da in unserem STD-Kennwert für Aggressions-Hemmung (AH) sowohl deren intrinsische als auch extrinsische Anteile konfundiert sind, ist ein Zusammenhang mit den Ratingvariablen GEWISS und TRAUTN gleichwahrscheinlich. Im Hinblick auf unsere differentiellen Skalen für Aggressions-Hemmung beziehen sich unsere Annahmen auf Zusammenhänge zwischen der Ratingvariablen GEWISS und der Testvariablen AHI einerseits und auf ebensolche zwischen der Ratingvariablen TRAUTN und der Testvariablen AHE andererseits. Aufgrund der bipolaren Beziehung zwischen den Ratingvariablen WEHREN und TRAUTN sind negative Zusammenhänge zwischen der Ratingvariablen WEHREN und unseren Testvariablen für Aggressions-Hemmung verständlich, vor allem mit den Meßwerten der DIFF-Skala AHE.

3.4.4 Methode

Die Datenerhebung erfolgt im Juli 1982 an fünf saarländischen Grundschulen. Wir beziehen alle dort eingerichteten 4. Klassen, insgesamt 12, in die Untersuchung ein. Sie findet während der normalen Unterrichtszeit statt. An der Untersuchung werden auch die Mädchen einer Klasse beteiligt; deren Daten gehen jedoch nicht in die Auswertung ein.

Die Ausgangsstichprobe umfaßt 128 Jungen. Durch Teilnahmeverweigerung, mangelnde Deutschkenntnisse (Lese- und Rechtschreibschwierigkeiten kombiniert oder auch nur erhebliche Leseschwierigkeiten,

zum Teil verbunden mit ausländischer Staatszugehörigkeit), Krankheit und anderes bedingt, geht deshalb nur eine Stichprobe von n = 77 Jungen in die Auswertung ein. Das entspricht einer Teilnahmequote von 50.2 %. Die Altersverteilung der Probanden ist in Tabelle 36 dargestellt.

	ALTER IN JAHREN			
HÄUFIGKEIT	9 Jahre	10 Jahre	11 Jahre	12 Jahre
absolut	3	61	11	2
prozentual	3.9	79.2	14.3	3.6

Tab. 36 Altersverteilung der Probanden der Untersuchung

Zunächst wird das AM-GITTER im Gruppenversuch vorgelegt. Anschließend werden die Ratings im Einzelversuch erhoben. An der Durchführung der Untersuchung sind drei Personen beteiligt.

Wir werten das AM-GITTER auf allen vorgeschlagenen Subskalen aus und benutzen die Rohwerte. Wir werten das Ratingverfahren auf allen sieben Skalen aus und benutzen ebenfalls die Rohwerte. Diese sind als das arithmetische Mittel aus den vier Einzelschätzungen (die jeweils zwischen 1 und 7 variieren können) bestimmt. Zur Bestimmung von Zusammenhängen in den Meßwertereihen berechnen wir SPEARMAN-Rangkorrelationen (bindungskorrigiert) und prüfen deren Signifikanz zweiseitig mit dem t-Test von FISHER (CLAUSS & EBNER 1975, S. 274 f.).

3.4.5 Ergebnisse

Das Ergebnis der Korrelationsanalyse ist in Tabelle 37 dargestellt.

rho	A	AH	n-A.	Ges.-A.	ADM	ADH	AHI	AHE
TEILEN	.05	-.06	.09	.00	.12	-.01	-.08	-.01
AUTAGG	-.08	.02	-.04	-.07	-.10	-.07	.11	-.11
TRAUTN	.01	-.11	.08	-.01	.02	-.08	-.13	-.04
HELFEN	.11	.00	.07	.10	.18	.03	-.05	.08
RAUFEN	-.13	.09	-.13	-.02	-.19	-.04	.17	-.05
GEWISS	-.08	.05	-.08	.06	-.11	-.12	.03	.09
WEHREN	-.10	.20 +	-.20 +	.03	-.12	.00	.27 ++	.09

+ : sig. 4 % / ++ : sig. 0.9 %

Tab. 37 Rangkorrelationen zwischen den Ratingvariablen und den einzelnen Subskalen des AM-GITTERS (n = 77; SI/SPSS, NONPAR CORR)

Es zeigen sich meist nichtsignifikante, um Null streuende Korrelationen. Die wenigen Ausnahmen sind nicht im Sinne unserer eingangs formulierten Zusammenhangsannahmen interpretierbar. Dazu zeigen sie sich in umgekehrter Richtung wie erwartet und gerade an der Stelle nicht, an der wir sie noch am ehesten für möglich gehalten hätten (WEHREN-AHE). Dennoch sind die Beziehungen zu dem Außenkriterium WEHREN unter einer bestimmten Voraussetzung insgesamt sinnvoll,

worauf wir jedoch erst in der Diskussion der Ergebnisse eingehen wollen. Zunächst jedoch geben die vorliegenden Ergebnisse zu Überlegungen Anlaß, die Daten im Hinblick auf einige der eingangs genannten kritischen Punkte (Reliabilität der Ratings, Dimensionalität und Bereichskongruenz der Außenkriterien) weiterführend zu analysieren, bevor wir die Analysen zur Korrelation mit Außenkriterien vorerst abschließen.

3.4.6 Weiterführende Analysen

3.4.6.1 Vorüberlegungen

Die Inspektion der Ratings ergibt, daß teilweise eine nur sehr geringe Übereinstimmung zwischen den Ratern hinsichtlich der Beurteilung eines Probanden in einem Außenkriterium besteht. Möglicherweise ist dieser reliabilitätsmindernde Faktor durch das Mittelungs-Schätzverfahren noch nicht hinreichend kompensiert. Dies führt zu der Überlegung, unter Beibehaltung der Berechnungsmethode für die Ratings die Zusammenhangsanalysen in Teilstichproben zu wiederholen, die nach dem Ausmaß der Interrater-Übereinstimmung kontinuierlich abgestuft sind.

Eine weitere Überlegung beruht auf den Ergebnissen zur Dimensionalität der Ratings. Diese lassen den Schluß zu, daß die einzelnen Außenkriterien mit unserer Methode unterschiedlich "quellnahe" d.h. mit unterschiedlicher Güte erfaßt werden könnten. Vorhandene vs. fehlende Wehrbereitschaft sind von den ins Auge gefaßten Kriterien

offensichtlich am unproblematischsten, intrinsische Aggressions-Hemmung dagegen am schwierigsten einer angemessenen Beurteilung zugänglich. Möglicherweise läßt sich deshalb der Vergleichsgegenstand zum Validitätsnachweis mittels einer Linearkombination mehrerer Außenkriterien präziser erfassen. Aufgrund der zwischen den Ratings bestehenden Korrelationsverhältnisse scheint jedoch auch die Auspartialisierung der jeweils nicht in Betracht stehenden aggressions- und hemmungsthematischen Außenkriterien methodisch sinnvoll. Dieser, die Außenkriterien etwas anders gewichtende Berechnungsansatz hatte ja bei der Vergleichsmessung PFT (vgl. Abschnitt 3.3.6.3.1) als einziger zu überzufälligen und dabei theoretisch vernünftigen Beziehungen geführt.

Eine dritte Überlegung führt schließlich zu der Frage, ob die beiden Erhebungsverfahren (Ratingmethode, AM-GITTER) überhaupt einen gemeinsamen Faktorenraum besitzen ("faktorielle Validität"). Die Frage nach der faktoriellen Validität des AM-GITTERs ist teilweise schon beantwortet, denn der Begriff beruht lediglich auf der besonderen Art des methodischen Zugangs, mit dem hier versucht wird, Validität nachzuweisen (vgl. SCHMIDTCHEN 1975, S. 23). Man faktorisiert dabei zum Teil ausschließlich Information, die das zu prüfende Testverfahren selbst beinhaltet, und insofern ist faktorielle Validität "ausschließlich eine Funktion der Iteminterkorrelationen" (DIETERICH 1973, S. 181), zum Teil aber auch Information auf der Grundlage mehrerer Erhebungsinstrumente zum gleichen Gültigkeitsbereich, denen das zu prüfende Testverfahren dann ebenfalls angehört; dabei erhält man eine besondere Form der Übereinstimmungsvalidität (concurrent validity).

Bei erhöhter Reliabilität der Ratings sollten sich bestehende Zusammenhänge in den Daten insgesamt deutlicher abzeichnen und damit auch diejenigen, die wir aufzeigen möchten. Bei Auspartialisierung nicht in Betracnt stehender Außenkriterien sollten sich die Voraussetzungen, die o.g. Beziehungen nachzuweisen, ebenfalls verbessern. Bei der Bildung von Linearkombinationen sowohl aus Ratings wie aus Testvariablen des AM-GITTERs sind kaum Voraussagen möglich, an welcher Nahtstelle der beiden mehrdimensionalen Datensätze die besten Voraussetzungen zum Aufzeigen signifikanter Zusammenhänge bestehen. Vergleichbares gilt für die gemeinsame Faktorisierung von Test- und Außenkriterien-Variablen.

3.4.6.2 Methode

Zur Verbesserung der Reliabilität der Ratings bilden wir pro Rating-Variable drei verschiedene Untergruppen von Probanden, und zwar eine Gruppe D_0 (völlige Übereinstimmung der Rater bezüglich des vergebenen Punktwertes), eine Gruppe D_1 (maximale Abweichung in den vergebenen vier Einschätzwerten des Rater-Teams gleich 1) und eine Gruppe D_2 (maximale Abweichung gleich 2). In diesen Untergruppierungen rechnen wir die Korrelationsanalysen in der bereits dargestellten Art und Weise noch einmal durch.

Zur Untersuchung auf Zusammenhänge zwischen Linearkombinationen aus Testvariablen des AM-GITTERs und Linearkombinationen aus den

aggressions- und hemmungsthematischen Außenkriterien führen wir eine kanonische Analyse (vgl. GAENSSLEN & SCHUBÖ 1973; MOOSBRUGGER 1978) durch. Wir geben in die Analyse die Rohwerte der Außenkriterien-Variablen AUTAGG, TRAUTN, RAUFEN, GEWISS und WEHREN ein. Als Testvariablen berücksichtigen wir die Rohwerte in den Subtests A, AH, ADM, ADH, AHI und AHE; daneben in zwei weiteren Variablen, nämlich die ausschließlich auf die Bildsituation Nr. 11 (Schule) bezogenen Rohwerte für das Aggressionsmotiv (A_{11}) und die Aggressions-Hemmung (AH_{11}). Die Variablen A_{11} und AH_{11} würden für den Fall die Kongruenz zwischen Rating-Variablen und AM-GITTER erhöhen, daß der Gesichtspunkt, ob die Gleichaltrigeninteraktionen innerhalb oder außerhalb der Schule stattfinden, bei den Ratings Relevanz besitzt. Die Signifikanz der kanonischen Korrelationen prüfen wir wiederum über die chi^2-Approximation von WILK's Lambda (vgl. MOOSBRUGGER 1978, S. 119 ff.).

Zur Eliminierung des Einflusses derjenigen Außenkriterien, auf die sich unser Korrelationsmaß im univariaten Fall nicht bezog, berechnen wir multiple partielle Korrelationen 4. Ordnung. Wir korrelieren die Testvariablen des AM-GITTERs mit den Außenkriterien AUTAGG, RAUFEN, WEHREN, TRAUTN und GEWISS (beidemal Rohwerte) einzeln unter jeweiliger Kontrolle der 4 restlichen Außenkriterien. Die Signifikanz der partiellen Korrelationen prüfen wir zweiseitig nach dem t-Test von FISHER (CLAUSS & EBNER 1975, S. 274 f.).

Zur Untersuchung der Meßwerte des AM-GITTERs gemeinsam mit den Außenkriterien im faktorenanalytischen Modell (THURSTONE 1969) benutzen wir die Rohwerte auf allen Subskalen des AM-GITTERs sowie die Rohwerte in allen Außenkriterien. Dazu berechnen wir im AM-GITTER einen Rohwert, dem wir entnehmen können, wie oft die freundschaftsthematischen Statements (F-Statements) über alle Bildsituationen hinweg mit JA beantwortet wurden und geben diesen Wert als Variable F ebenfalls in die Analyse ein.

3.4.6.3 Ergebnisse

RELIABILITÄTSKONTROLLE: Auch in den neuen Untergruppierungen mit nun recht hoher Interrater-Übereinstimmung ergeben sich keine substantiellen Änderungen in den Korrelationen, sodaß auf deren Darstellung hier verzichtet werden kann.

PARTIELLE KORRELATIONEN: Die partiellen Korrelationen führen insgesamt nicht zu neuen Erkenntnissen. Sie liefern bis in Einzelheiten das bereits aus Tabelle 37 bekannte Bild, sodaß wir auf deren Darstellung hier ebenfalls verzichten.

KANONISCHE KORRELATIONEN: Das Ergebnis der kanonischen Analyse ist in Tabelle 38 dargestellt. Diesem zufolge läßt sich keine signifikante kanonische Korrelation zwischen einer Linearkombination aus Außenkriterien für Aggression und Aggressions-Hemmung und einer Linearkombination aus Testvariablen des AM-GITTERs aufzeigen. Damit besteht auch kein übergreifender Zusammenhang.

1. kanonische Korrelation	=	0.44
Eigenwert	=	0.19
WILK´s Lambda	=	0.58
chi^2	=	38.06
df	=	40.00
Signifikanzniveau	=	0.56

Tab. 38 Ergebnis der kanonischen Analyse ausgewählter Testvariablen des AM-GITTERs und Außenkriterien (n = 77; SI/SPSS, CANCORR)

FAKTORIELLE ANALYSE: Der Scree-Test verweist deutlich auf vier extraktionswürdige Faktoren. Die betrachteten Vierfaktorenlösungen weisen jedoch unzureichende Kommunalitäten für die Subtestvariablen des AM-GITTERs auf. Wir begnügen uns deshalb mit der Extraktion der ersten drei Faktoren und erhalten nach genau einer Iteration die in Tabelle 39 dargestellte, unrotierte Ladungsmatrix. Im Gegensatz zu den Vierfaktorenlösungen können in der vorliegenden Dreifaktorenlösung nicht alle Rating-Variablen dimensional befriedigend aufgeklärt werden, da die Rating-Dimension PROSOZIALES VERHALTEN erst mit der Extraktion des 4. Faktors zum Vorschein kommt. Die Dimensionen der unrotierten Dreifaktorenlösung stellen sich wie folgt dar: Mit Faktor 1 erfassen wir die bereits aus der Entwicklung des AM-GITTERs bekannte Dimension SOZIALE WERTUNG. Faktor 2 stellt den zweiten GITTER-Faktor, die thematische Dimension AGGRESSION dar. Faktor 3 schließlich ist die aus der Strukturanalyse der Ratings bekannte zweite bipolare Dimension, nämlich WEHRBEREITSCHAFT vs. WEHRANGST. Im Hinblick auf faktorielle Invarianz der internen Struktur des AM-GITTERs liegt damit ein wichtiges

Nebenergebnis vor, das wir um folgende Angaben ergänzen wollen: Nach VARIMAX-Rotation (die auch hier etwa 1/8 Volldrehung des unrotierten Systems entgegen dem Uhrzeigersinn entspricht) erhalten wir eine Struktur mit den Faktoren AGGRESSIONSMOTIV (F2), AGGRESSIONS-HEMMUNG (F1) und WEHRBEREITSCHAFT vs. WEHRANGST (F3).

	UNROTIERTE FAKTOREN			
VARIABLE	Faktor 1	Faktor 2	Faktor 3	h^2
A	-.66	.72	.20	.99
AH	.79	.61	-.06	.99
n-A.	-.98	-.04	.16	.99
Ges.-A.	.23	.97	.09	.99
ADM	-.67	.66	.15	.91
ADH	-.60	.64	.28	.85
AHI	.85	.42	-.03	.90
AHE	.60	.70	-.12	.86
F	.68	-.08	-.12	.48
TEILEN	-.13	.13	-.13	.05
AUTAGG	.19	-.14	.38	.20
TRAUTN	-.27	.07	-.64	.49
HELFEN	-.07	.21	-.11	.06
RAUFEN	.28	-.19	.72	.63
GEWISS	.03	.05	-.39	.16
WEHREN	.38	-.10	.78	.76
Summe a_j^2	4.78	3.49	2.07	
% gemeinsame Varianz :	46.23	33.75	20.02	
% Gesamt-Varianz :	29.88	21.81	12.94	SUMME: 64.63

Tab. 39 Ergebnis der faktoriellen Analyse der Testvariablen des AM-GITTERs gemeinsam mit Außenkriterien (n = 77; SI/SPSS, FACTOR (PA 2))

3.4.7 Diskussion

Die vorliegenden Ergebnisse entsprechen - vordergründig betrachtet - nicht unseren eingangs formulierten Erwartungen. Denen zufolge hätte sich zumindest einmal für jedes der beiden Teilmotive eine substantielle Korrelation in der erwarteten Richtung mit einem dafür vorgesehenen Außenkriterium aufzeigen lassen müssen. Daran gemessen haben auch die Ergebnisse unserer weiterführenden Analysen an der Befundlage nichts Wesentliches geändert. Dies führt zwangsläufig zu der Frage, warum das so ist. Wir können, bevor wir näher auf die Einzelheiten zu sprechen kommen, das Ergebnis unserer Überlegungen an dieser Stelle bereits in Kürze zusammenfassen. Es zeigt sich, daß ein Validitätsnachweis durch Korrelation mit Außenkriterien offensichtlich schwieriger zu führen ist, als wir uns das zunächst vorgestellt hatten. Dafür gibt es theoretische Argumente als auch Hinweise in unseren eigenen Daten. Sollte dies zutreffen, dann handelt es sich bei der hier durchgeführten Studie im Grunde um eine Voruntersuchung. Die Ergebnisse einer unter verbesserten Bedingungen durchgeführten Nachfolge-Studie könnten definitiv zur Klärung der Frage beitragen, welcher Stellenwert den oben angesprochenen Fakten und Argumenten beizumessen ist. Es besteht aber auch die Möglichkeit, daß die Ergebnisse gerade so, wie sie ausgefallen sind, zu erwarten waren, worauf wir in unseren Vorüberlegungen hingewiesen haben (datentypologischer Ansatz von McCLELLAND). Wir werden sie im nächsten und letzten Kapitel dieser Arbeit auf ihren Realitätsgehalt prüfen, soweit das unsere Daten

zulassen. Nicht zuletzt, und das könnte der Realität durchaus am nächsten kommen, ist die Annahme wahrscheinlich, daß mehrere derjenigen Bedingungen, auf die wir nachfolgend zu sprechen kommen, gleichzeitig (in unbekannter Kombination und Gewichtung) zutreffen.

Die ersten Überlegungen, die wir im Hinblick auf die vorliegende Ergebnislage anstellen wollen, haben die Validität der Außenkriterien selbst zum Gegenstand. Hierbei ist die Annahme plausibel, daß Verteidigungsaggression zuverlässiger und daher mit einer besseren Voraussetzung für Validität erfaßt wird als feindselige Aggression und die dazugehörige Hemmung. Einen Hinweis darauf gibt die Anzahl der Probanden pro Rating-Variable, bei der die Rater um nicht mehr als zwei Punktwerte in ihren Einschätzurteilen differieren. Dies ist bei WEHREN in 42, bei AUTAGG in 28 und bei GEWISS in ebenfalls 28 Fällen von jeweils 77 zur Beurteilung anstehenden Probanden gegeben. Für TRAUTN - als Gegenpol zu WEHREN - errechnen wir 41 und für RAUFEN ebenfalls 41 Fälle. Aus den Dimensionierungsbefunden wissen wir jedoch, daß die Übereinstimmung der Rater in der Rating-Variablen RAUFEN mit darauf zurückgehen kann, daß sie Anteile einer habitualisierten Wehrbereitschaft beinhaltet. Nimmt man faktorielle Einfachstruktur und Reliabilität zusammen, so weist unter den Aggressionskriterien WEHREN die besten Voraussetzungen für Validität auf. In der Tat finden wir dort auch die einzigen signifikanten Korrelationen zwischen Testvariablen des AM-GITTERs und Außenkriterien, auf deren Interpretation wir an dieser Stelle der Diskussion noch nicht eingehen wollen.

Weitere Argumente für die bestehende Ergebnislage resultieren aus der Annahme, daß die Meßwerte des AM-GITTERs insgesamt kaum Extensitäts-, dagegen in hohem Maße Intensitätscharakter für einen eng umschriebenen Bereich möglicher Motivzentrierung besitzen. Auch hier ergeben sich wiederum ähnliche Schlußfolgerungen, was die Eignung der Außenkriterien zum Validitätsnachweis betrifft. Das Außenkriterium WEHREN ist dann gegenüber dem Außenkriterium AUTAGG zusätzlich noch im Hinblick auf die o.g. Bereichsklasse günstig, da kongruent.

Bei den Analysen zur Meßwertestabilität (vgl. Abschnitt 3.2) haben wir in einem Fall gezeigt, daß sich in einer nicht signifikant von Null verschiedenen Korrelation durchaus signifikante, jedoch entgegengerichtete Zusammenhänge überlagern können. Diese Besonderheiten lassen sich jedoch - falls existent - erst ausfindig machen, wenn man nach theoretisch sinnvollen Erwägungen Subgruppen gebildet und die Korrelationsanalyse dort noch einmal getrennt durchgeführt hat. Auf die hier zur Diskussion stehende Ergebnislage übertragen läuft das auf die Annahme einer differentiellen Validität der Meßwerte des AM-GITTERs in dem Sinne hinaus, daß sie in nach theoretischen Gesichtspunkten wohlunterschiedenen Populationen mit Außenkriterien unterschiedlich gerichtete Zusammenhänge aufweisen, die sich bei Nichtberücksichtigung der moderierenden Bedingungen überlagern. Sie sei an einer Untersuchung von LESSER (1957, 1959 a) verdeutlicht. LESSER erfaßte - in einer vergleichbaren Population - die Persönlichkeitsvariable Aggressivität mittels des (operanten) L-TAT (vgl. Abschnitt 3.3) und das zugehörige Außenkriterium via Einschätzungen

durch Gleichaltrige (durch eine guess who-Technik erhobene Ratings). Es lagen also vergleichbare Bedingungen wie in unserem Fall vor. Die beiden Meßwertereihen korrelierten nicht ($r_{FA,OA}$ = .07). LESSER brachte nun - theoriegeleitet - eine Moderatorvariable (Einstellungen und Praktiken der Mutter hinsichtlich der Aggression ihres Kindes) zum Einsatz und erzeugte auf dieser Dimension zwei Subgruppen. In der einen, in der die Mütter die Söhne zu Aggression ermutigten, zeigte sich nun eine signifikant positive Korrelation zwischen den beiden Meßwertereihen ($r_{FA,OA}$ = .43); in der anderen dagegen, in der die Aggression der Jungen durch die Mütter gegenteilig beeinflußt wurde, eine negative ($r_{FA,OA}$ = -.41, sig. 10 %). Beide Koeffizienten sind darüber hinaus signifikant voneinander verschieden. Die Hypothese differentieller Validität der Meßwerte des AM-GITTERs kann jedoch auf Grundlage der vorliegenden Daten nicht geprüft und sollte ggfls. in einer Nachfolgeuntersuchung geklärt werden.

Der in Tabelle 35 dargestellten Ladungsmatrix der Rating-Skalen ist zu entnehmen, daß wir (erfreulicherweise) zwischen feindseliger und instrumenteller Aggression in den Außenkriterien trennen. LESSER (1959 b) hat nun in einer der unseren vergleichbaren Population gefunden, daß physische Aggression auf Provokation hin als einzige der untersuchten Handlungsformen positiv mit einem Popularitätsindex korreliert war (ebd., S. 23). Auf unsere Ratings übertragen könnte das bedeuten, daß durch die aufgezeigte dimensionale Struktur der Ratingdaten (zusätzlich) eine

Dimension der SOZIALEN WERTUNG verläuft, die feindselige Aggression (und fehlende Wehrbereitschaft) in einer entgegengesetzten Richtung lädt wie Wehrbereitschaft (und prosoziales Verhalten); diese Vorstellung ist mit der o.g. Dreikomponentenstruktur durchaus vereinbar. In diesem Sinne wären dann auch die Ladungen der Variablen WEHREN und TRAUTN auf der Dimension SOZIALE WERTUNG in Tabelle 39 zu interpretieren, die dann (allerdings schwach ausgeprägt) für faktorielle Validität aller Meßwerte des AM-GITTERs sprächen. Über diese Hilfskonstruktion ließen sich dann natürlich auch die in Tabelle 37 aufgezeigten, signifikanten Korrelationen zu dem Außenkriterium WEHREN erklären. Insbesondere die intrinsischen Bestandteile der Aggressions-Hemmung unterliegen demnach einer gleichgerichteten sozialen Wertung wie eine habitualisierte Wehrbereitschaft, das Aggressionsmotiv selbst dagegen (und dort insbesondere wiederum die Bedürfniszentrierten Bestandteile) einer entgegengesetzten. Demnach beruhen diese Korrelationen gerade nicht auf thematischen Gemeinsamkeiten, was bei Definition des AM-GITTER-Faktors AGGRESSION als thematische Kategorie des Aggressionsmotivs natürlich zu fordern ist. Im Rahmen dieses Interpretationsansatzes sind auch die nichtsignifikanten Korrelationen der Testvariablen des AM-GITTERs mit dem Außenkriterium WEHREN theoretisch sinnvoll, da sie anhand der Zeigerstellung eines Subtestvektors im Generalfaktormodell vorausgesagt werden können, wenn die Korrelationskoeffizienten der Subtestvariablen n-A. und AHI mit WEHREN feststehen. Dies verweist auf einen wichtigen Nebenbefund der hier durchgeführten Analyse. Wir können

nämlich nun erstmals nach Abschluß der Entwicklung des AM-GITTERs zeigen, daß die Dimensionalität unserer Motivmeßwerte den konstruktiven Zielsetzungen entspricht (vgl. Tab. 39).

Wir haben nun eine Reihe von Überlegungen angestellt, die alle davon ausgehen, daß der Datentyp (sensu McCLELLAND) keine Rolle spielt, wenn Meßwerte korreliert werden. Insbesondere die LESSER-schen Befunde sind diesbezüglich interessant, da hier operante (L-TAT) und respondente Meßwerte (Ratings) miteinander korreliert werden. Nach McCLELLAND hätte man es bei der aufgezeigten Nullkorrelation bewenden lassen müssen, da sie nach seiner Theorie zu erwarten war. Demgegenüber stellt das Konzept der differentiellen Validität einen völlig anders gearteten Erklärungsansatz dar, der jedoch nicht im Sinne einer den ersten ausschließenden Alternative verstanden werden sollte. Wir waren ursprünglich durch die positiven Ergebnisse, die OLWEUS mit dieser Art von Außenkriterium berichtete, darauf gekommen, nach diesem Ansatz zu verfahren. Wir hatten jedoch zu wenig gesehen, daß bei OLWEUS die Aggressivität mit Fragebogen gemessen wurde und somit (nach McCLELLAND) datentypologische Kompatibilität bestand.

Nach McCLELLAND ist also die Ergebnislage genau so zu erwarten gewesen, falls das AM-GITTER TOT-Daten liefert. Wir verfügen nun über Voraussetzungen, diese Argumentation zu prüfen, indem wir - wie eingangs erwähnt - das Verhalten der Rater beim Einschätzprozeß als Außenkriterium wählen und versuchen, es als spezifisches Attributionsverhalten unter aggressionsmotivationstheoretischen Gesichtspunkten zu erklären. Diesem Anliegen wenden wir uns im nächsten Kapitel zu.

3.5 Attribution

3.5.1 Vorüberlegungen

Bei der nachfolgend durchgeführten Untersuchung handelt es sich formal um eine ex post facto - Analyse, die zunächst als Validierungsbeitrag nicht vorgesehen war. Wir sehen uns jedoch aufgrund der besonderen Ergebnislage, die wir bei dem Versuch, Validität durch Korrelation mit Außenkriterien aufzuzeigen, zu Tage gefördert haben, zu diesem Schritt veranlaßt. Nach dem o.g. Validierungsansatz bestehen schwache Hinweise auf Korrelation mit Außenkriterien, die darauf zurückzuführen ist, daß Meßwerte des AM-GITTERs wie Außenkriterien den Faktor SOZIALE WERTUNG laden. Dagegen haben wir keinen Hinweis auf Korrelation mit Außenkriterien gefunden, der auf einen gemeinsamen Faktor FEINDSELIGE AGGRESSION verweist. Wir haben darauf versucht, Gründe für diese Befundlage zu finden. Sie kann zunächst darauf beruhen, daß unser Verfahren zur Erhebung der Außenkriterien erst noch weiterentwickelt werden muß, bevor es hinreichend valide Information liefert. Sie kann im weiteren darauf beruhen, daß unser Motivtest differentiell valide ist und der experimentelle Ansatz zum Validitätsnachweis daher im Grunde unzulänglich. Es ist nun keine Frage, daß sowohl die Weiterentwicklung unseres Erhebungsverfahrens für Außenkriterien als auch die Prüfung des AM-GITTERs auf differentielle Validität angezeigte Maßnahmen sind, da der o.g. Korrelationsansatz aller Voraussicht nach theoretisch differenziertere Erwägungen erfordert, als wir sie hier vorausgesetzt haben. Es ist jedoch nicht mehr Gegenstand dieser Arbeit,

das zu leisten. Als einziges verbleibt uns noch die Prüfung eines dritten Arguments für die Ergebnislage der Korrelationsstudie mit Außenkriterien, das sich aus dem datentheoretischen Ansatz von McCLELLAND (vgl. Abschnitte 3.3.1 und 3.4.1) herleiten läßt. Diesem zufolge sind erst Korrelationen zwischen Test- und Verhaltensdaten zu erwarten, wenn beide im Datentyp übereinstimmen. Unter der Voraussetzung, daß unser Testverfahren Daten vom operanten Typus liefert, wäre somit Übereinstimmung gegeben, wenn wir das Urteil der Rater als charakteristische Verhaltenseigentümlichkeit dieser selbst betrachteten und nicht als festgestellte Verhaltenseigentümlichkeit derjenigen, auf die es sich bezieht. Damit würde das Beurteilungsverhalten selbst zum Außenkriterium und auf Zusammenhänge zum Aggressionsmotiv und der Aggressions-Hemmung der Urteiler (Rater) zu untersuchen sein. Genau das wollen wir nun prüfen.

Aufgrund des bereits erwähnten ex post facto - Charakters der nachfolgend dargestellten Untersuchung erübrigt sich die Formulierung spezieller Hypothesen, was Art und Richtung der Zusammenhänge angeht als auch den Ort, an dem wir sie finden. Wir gehen statt dessen lediglich der Frage nach, ob sich die Befundlage unserer Korrelationsstudie mit Außenkriterien ändert, wenn wir die Rating-Daten unter einem neu formulierten Gesichtspunkt inspizieren. Damit meinen wir, daß sich nun deutlicher als bisher signifikante Beziehungen zwischen Test- kriterialen Verhaltensdaten aufzeigen lassen.

3.5.2 Fragestellung

Es besteht nun die Möglichkeit, das Einschätzverhalten der Rater über die gesamte Bandbreite der zuzuschreibenden Verhaltensmerkmale zu untersuchen. Wir halten dies aufgrund der bereits zu den einzelnen Rating-Skalen geäußerten Kritik als auch im Hinblick darauf, die Ergebnisse vergleichsweise einfach und eindeutig interpretieren zu können jedoch für weniger angebracht. Statt dessen beschränken wir uns auf die Auswahl zweier Verhaltensmerkmale, die im o.g. Sinne günstig scheinen.

Als erstes wählen wir das Verhaltensmerkmal AUTAGG. Es wird gänzlich durch die Dimension FEINDSELIGE AGGRESSION beschrieben. Im Gegensatz dazu ist das Verhaltensmerkmal RAUFEN als faktoriell komplex zu betrachten und eignet sich daher zu einer Untersuchung des Attributionsverhaltens weniger gut. Die etwas geringere Übereinstimmung der Rater in der Beurteilung der Ratees hinsichtlich AUTAGG im Vergleich zu RAUFEN (vgl. 3.4.7) ist hier eher von Vorteil, da wir ja die individuellen Abweichungen der Rater in ihrem Urteil über die Ratees gerade mit zum Gegenstand unserer Analyse machen. Bei maximaler Interrater-Übereinstimmung wäre diese Quelle der Rating-Varianz versiegt, und somit wären die Ratings auf Rater-Ebene nicht mehr analysierbar.

Als zweites wählen wir das Verhaltensmerkmal WEHREN. Dieses Merkmal wird ebenfalls gänzlich durch eine aggressionsthematische Dimension beschrieben, die jedoch - und das läßt einen interessanten

Vergleich zu - gerade nicht die Zuschreibung feindseliger Aggression und Verletzungsabsicht beinhaltet.

Unsere Fragestellung läßt sich nun dahingehend konkretisieren, Unterschiede im Attributionsverhalten der Rater hinsichtlich der o.g. Verhaltensmerkmale in Zusammenhang mit individuellen Differenzen in ihren überdauernden Persönlichkeitsmerkmalen Aggressionsmotiv und Aggressions-Hemmung zu bringen, um sie ggfls. im Sinne motivbedingter Attributionseffekte interpretieren zu können.

3.5.3 Methode

Zur Durchführung der Analyse stehen die Daten von 48 Ratern zur Verfügung. Wir wählen daraus die Daten von 36 Ratern zur weiteren Betrachtung aus. Die nicht in die Analyse einbezogenen Rater gehören Klassen mit zu geringer Anzahl von Jungen (Ratees) an, sodaß sie nach unserem methodischen Ansatz keine verwertbare Information mehr erbringen.

Bei der Operationalisierung des Attributionsverhaltens der Rater gehen wir folgendermaßen vor: Pro Verhaltensmerkmal und Klasse getrennt berechnen wir (a) wie hoch alle Ratees im Durchschnitt von allen Ratern eingeschätzt werden (= Rater-TEAM-$\bar{x}$) und (b) wie hoch alle Ratees im Durchschnitt individuell von jedem einzelnen Rater eingeschätzt werden (= RATER-$\bar{x}$). Zur Bildung des Durchschnitts wählen wir das arithmetische Mittel. Durch Vergleich von RATER- und

TEAM-$\bar{x}$ läßt sich nun feststellen, ob ein Rater bezüglich der Attribution eines Verhaltensmerkmals im allgemeinen nach oben oder unten von der durchschnittlichen Einschätzhöhe abweicht. Die Berechnung von RATER- und TEAM-$\bar{x}$ erfolgt ausschließlich auf Grundlage von Probanden, die dem Rater-Team selbst nicht angehören. Da wir auf klassenspezifische Einschätznormen als Vergleichsmaßstab für die individuellen Rater-Abweichungen rekurrieren, gestattet uns diese Vorkehrung eine angemessene Beurteilung des Rater-Verhaltens in der zusammenfassenden Betrachtung über verschiedene Klassen hinweg.

Die überdauernden Persönlichkeitsmerkmale der Rater erfassen wir mit den acht Subskalen des AM-GITTERs. Dabei geben wir in die Analyse jeweils die Rohwerte ein.

Zur Aufdeckung von Zusammenhängen zwischen Persönlichkeitsmerkmalen und Attributionsverhalten der Rater bilden wir pro Skala des AM-GITTERs jeweils zwei Extremgruppen, hoch- und niedrigmotivierte Rater. Kriterium für die Zugehörigkeit dazu ist, ob der Rater dem Bereich unter dem 1. oder über dem 3. Quartil der Meßwerteverteilung aller Rater in dieser Testvariablen angehört. Daraus resultieren einheitlich Besetzungszahlen von n=9 pro Extremgruppe. Anschließend zählen wir aus, wie viele Rater pro Extremgruppe nach oben beziehungsweise nach unten von ihrem zugehörigen TEAM-$\bar{x}$ abweichen. Das Ergebnis tragen wir in ein Vierfelderschema ein. Den Zusammenhang kennzeichnen wir über ein Vierfelder-phi, das bei unserer speziellen Anordnung der Ergebnisse

auch richtungsspezifisch interpretiert werden darf. Die Signifikanz des Zusammenhangs prüfen wir über die chi^2 -Verteilung (vgl. CLAUSS & EBNER 1975, S. 281 f.).

3.5.4 Ergebnisse

Die Ergebnisse unserer Analyse sind im einzelnen der Tabelle 40 zu entnehmen.

phi	AUTAGG	WEHREN
AHI	-.47 +	.00
AH	-.34	-.11
AHE	-.33	.11
Ges.-A.	.00	-.11
ADH	.45	-.22
A	.56 ++	-.45
ADM	.56 ++	-.45
n-A.	.67 +++	-.33

\+ : sig. 5 %
++ : " 2.5 %
+++ : " 1 %

Tab. 40 Zusammenhänge zwischen Testvariablen des AM-GITTERs und Attribution feindseliger Aggression bzw. Wehrbereitschaft (n = jeweils 18)

Es zeigen sich nun deutlich systematische Attributionsfehler im Zusammenhang mit speziellen Persönlichkeitsvariablen aus dem Bereich Aggressionsmotiv als auch aus dem Bereich Aggressions-Hemmung. Diese Attributionsfehler können wir hinsichtlich der Zuschreibung von Aggression gegen Autoritätspersonen in feindseliger Absicht (AUTAGG) zufallskritisch absichern. Demzufolge neigen Jungen mit hohem Aggressionsmotiv zu einem systematischen Attributionsfehler im Sinne einer erhöhten Zuschreibung des Verhaltensmerkmals AUTAGG, Jungen mit einem niedrigen Aggressionsmotiv zu einem systematischen Attributionsfehler im Sinne einer verminderten Zuschreibung des gleichen Verhaltensmerkmals. Jungen mit hoher Aggressionshemmung dagegen neigen zu einem systematischen Attributionsfehler im Sinne einer verminderten Zuschreibung, Jungen mit niedriger Aggressions-Hemmung zu einem systematischen Attributionsfehler im Sinne einer erhöhten Zuschreibung des Verhaltensmerkmals AUTAGG. Hinsichtlich der Attribution (einer Bereitschaft zu) reaktiver Verteidigungsaggression (WEHREN) können wir vergleichbare Effekte nicht zufallskritisch absichern. Es deuten sich jedoch auch hier systematische Attributionsfehler dergestalt an (die auf die Persönlichkeitsvariable Aggressionsmotiv beschränkt bleiben und bei einer geringfügigen Änderung in der Besetzungszahl einer Extremgruppe ($n = 10$) auf 2.5 % - Niveau signifikant werden; A: phi = -.50 ; ADM: phi = -.50), daß Jungen mit hohem Aggressionsmotiv zu einer verminderten, Jungen mit niedrigem Aggressionsmotiv zu einer erhöhten Zuschreibung des Verhaltensmerkmals WEHREN neigen.

3.5.5 Diskussion

Neuere empirische Untersuchungen (DODGE 1980, 1981, 1982; DODGE & FRAME 1982; vgl. PARKE & SLABY 1983, S. 571 ff.) zeigen, daß aggressive Jungen zu einem systematischen Attributionsfehler neigen, dem zufolge sie dem Handlungspartner vor allem in sozialen Situationen, die mehrere Deutungsmöglichkeiten zulassen, bevorzugt feindselige Absichten unterstellen. Im Gegensatz neigen nichtaggressive Jungen dazu, unter den gleichen Bedingungen eher freundliche Handlungsabsichten zu attribuieren.

Auch für die Formulierung eines Zusammenhangs in umgekehrter Richtung, daß nämlich die "Wahrnehmung" bzw. Zuschreibung von Verletzungsabsicht eine wichtige mediatorische Funktion bei der Steuerung und Intensitätsdosierung der eigenen Aggression gegenüber dem Handlungspartner besitzt, liegen stützende empirische Befunde vor (vgl. DENGERINK 1976).

Damit bestehen nun deutliche Hinweise für die Annahme der Verhaltenswirksamkeit auch der von uns aufgedeckten Attribuierungstendenzen hoch und niedrig aggressionsmotivierter sowie hoch und niedrig aggressionsgehemmter Rater hinsichtlich der Initiation und weiteren Gestaltung von Aggression. Diese Annahme könnte man - im nächsten Schritt - im Rahmen eines theoriegeleiteten Experiments einer direkten Prüfung zugänglich machen, um die Validierung insgesamt nun mehr auf Konstrukt-Ebene fortzusetzen.

Die systematischen Attributionsfehler der Rater hinsichtlich der

Zuschreibung absichtlich herbeigeführter, feindseliger Aggression gegen den Lehrer belegen in diesem Zusammenhang deutlich die Konstruktvalidität der Meßwerte des AM-GITTERs. Diese Annahme wird insbesondere dadurch gestützt, daß diese, für Aggressive bzw. Nichtaggressive typische Abweichungssystematik in der Attribution nicht mehr auftritt bzw. sich ändert, wenn die Aufgabenstellung die Attribution eines Verhaltensmerkmals erfordert, das sich gerade von dem oben genannten durch das Nichtverknüpftsein mit einer feindseligen Handlungsintention unterscheidet (Verhaltensmerkmal WEHREN).

Im Gegensatz zu den experimentellen Bedingungen, auf denen die Ergebnisse von DODGE beruhen, liegt in unserem Fall jedoch weder eine hypothetische noch eine tatsächliche Involviertheit der Rater in eine handlungsmäßige Bewältigung einer sozialen Situation vor, an der die Eingeschätzten (Ratees) als Handlungspartner unmittelbar beteiligt sind.

Bedenkt man also die distanzierte und entspannte Situation, in der die Ratings erhoben wurden, so ist es naheliegend, in den Attributionsfehlern die Auswirkung systematischer Effekte anzunehmen, die auf funktionellen Charakteristika des überdauernden Aggressionsmotivs und der überdauernden Aggressions-Hemmung beruhen (Motiv-Effekte). Speziell in diesem Zusammenhang von Interesse ist, daß es gerade die Bedürfnis-bezogenen Anteile des Aggressionsmotivs (gemessen mit der DIFF-Skala ADM) und die intrinsischen Anteile der Aggressions-Hemmung (gemessen mit der DIFF-Skala AHI) sind, auf die wir

die oben genannten Effekte isolieren können (vgl. Tab. 40).

Um die Zusammenhänge zwischen Diskriminationsleistung und Dimensionalität der einzelnen Subskalen des AM-GITTERs besser verstehen zu lernen, haben wir letztere in Tabelle 40 von oben nach unten analog der Abfolge ihrer Testvektoren im Generalfaktormodell (vgl. Abb. 14, Abschnitt 2.8.3, Teil II) aufgelistet. Wir können nun an den phi-Koeffizienten ablesen, wie sich Attributionsfehler und Richtung des Fehlers ändern, wenn wir Messungen in unterschiedlichen Bereichen des Aggressionsmotiv-Systems vornehmen. Messen wir in Bereichen, die einer ausgesprochen deutlichen sozialen Ablehnung unterliegen, kommt es zwischen Hoch- und Niedrigmotivierten zu einer gerade entgegengesetzten Abweichungssystematik in der Attribution, als wenn wir in Bereichen messen, die sich deutlich durch soziale Billigung und Befürwortung kennzeichnen lassen. Auch dies ist motivationstheoretisch sinnvoll, d.h. aus den unterschiedlichen Handlungszielperspektiven der Konstrukta Aggressionsmotiv und Aggressions-Hemmung heraus ableitbar und belegt die Eigenständigkeit unserer dafür vorgesehenen Maße. Auch die Triftigkeit der Einbeziehung der Aggressions-Hemmung in den abgeleiteten Kennwert n-A. wird in unseren Daten deutlich, da wir mit letzterem offensichtlich mehr attributionsrelevante Information erfassen, als mit dem Kennwert für das Aggressionsmotiv (A) allein. Nicht zuletzt sind es die differentiellen Skalen, die uns einen Einblick in die Zusammenhänge zwischen Dimensionalität einer Subskala (Zeigerstellung ihres Testvektors) und Diskriminationsleistung hinsichtlich

der hier aufgezeigten aggressionsrelevanten Attributionsmerkmale der Probanden gestatten. Diese Zusammenhänge sehen im vorliegenden Fall etwa so aus, daß die Diskriminationsleistung einer Subskala dadurch bestimmt zu werden scheint, in welchem Verhältnis ihre Kommunalität auf Korrelationsanteilen mit der Dimension SOZIALE WERTUNG beruht. Idealisiert können wir das so formulieren: Wandert der Testvektor - bei gleichbleibender Kommunalität - aus der Mittenposition (kollinear mit dem Faktor AGGRESSION) zu den Außenpositionen (kollinear mit dem Faktor SOZIALE WERTUNG) hin, steigt die Diskriminationsleistung kontinuierlich an.

Zusammengefaßt sprechen also die aufgezeigten Attribuierungsfehler, von denen wir annehmen dürfen, daß sie für aggressive und nichtaggressive Jungen in der vorgefundenen Art und Weise charakteristisch sind und sich unmittelbar auf deren Aggressionshandeln auswirken, für die Konstruktvalidität unserer Meßwerte. Darüber hinaus stützt diese Befundlage die Datentheorie von McCLELLAND, die wir mit in die Vorüberlegungen zur Durchführung dieser Analyse einbezogen haben.

3.6 Zusammenfassung

Wir haben nun die Validität der Meßwerte des AM-GITTERs erstmals unter diversen Gesichtspunkten geprüft. Wenn wir die dabei eruierten Befunde kritisch zusammenfassen, so zeichnet sich insgesamt etwa folgendes Bild ab:

MESSWERTESTABILITÄT: Auf der Ebene individueller Kennwerte erreichen wir nach einem Meßzeitintervall von 8 Wochen mit unserem Instrument - relativ unabhängig von der Skala - eine Retestkorrelation von etwa .71. Damit liegt - in Ergänzung zu den Konsistenzanalysen - eine Schätzung für die untere Grenze der Meßgenauigkeit vor. Vergleiche mit anderen Persönlichkeitsmeßverfahren für Kinder zeigen, daß wir uns damit im Bereich der Norm bewegen. Dabei besteht jedoch ein Zusammenhang zwischen der Höhe der Retestkorrelation und der Größe des Meßzeitintervalls. Konventionellerweise nimmt man nun an, daß sich während eines kurzen Meßzeitintervalls das zu erfassende Persönlichkeitsmerkmal weniger verändert und somit höhere Koeffizienten erwartet werden können als bei einem längeren (vgl. z.B. RAUCHFLEISCH 1980, S. 53). Demnach wäre ein negativer Zusammenhang zwischen Höhe der Retestkorrelation und Dauer des Meßzeitintervalls anzunehmen. Unsere Befundlage, die auf zusätzlichen Meßwiederholungen nach Intervallen von 2 und 4 Wochen beruht, widerspricht dieser Vorstellung. Interessanterweise verzeichnen wir in der Entwicklung der Meßwertestabilitäten über unterschiedliche Meßzeitintervalle hinweg bei dem 4 Wochen - Intervall einen leichten "Knick". D.h.: Gegenüber dem

2 Wochen - Intervall nimmt die Höhe der Retestkorrelation bei dem 4 Wochen - Intervall ab und steigt bei einer Meßwiederholung nach 8 Wochen wieder an. Da dieser Entwicklungsverlauf uns auch charakteristisch für die Stabilitätskoeffizienten des LM-GITTERs (SCHMALT 1974) zu sein scheint, vermuten wir in dem o.g. "4-Wochen-Knick" einen spezifischen Effekt.

Auf der Ebene Gruppen-bezogener Kennwerte verzeichnen wir bei einer Meßwiederholung nach 8 Wochen stabile Meßwerteverteilungen auf allen Subskalen des AM-GITTERs. Wir verfügen demzufolge nicht nur zur Kennzeichnung von Individuen (idiographischer Verwendungszweck), sondern auch zur Kennzeichnung von Gruppen (nomothetischer Verwendungszweck) über ein Verfahren, das man als zuverlässig betrachten kann.

DISKRIMINANTE UND KONVERGENTE VALIDITÄT: Theoretische Überlegungen ließen die Ausgangsbedingungen zum Nachweis konvergenter Validität unserer Meßwerte als schwierig erscheinen. Wir unternahmen dennoch diesen Versuch, mit auch, um Informationen darüber zu erhalten, in wieweit die Meßwerte unserer Skalen mit denjenigen zum gleichen Gegenstandsbereich eingeführter Skalen korreliert sind. Die Ergebnisse zeigen, daß wir ein eigenständiges Meßverfahren für kindliche Aggressivität entwickelt haben, das zu den zum Vergleich herangezogenen aggressionsrelevanten Subskalen des PFK (SEITZ & RAUSCHE 1976) keine und zu den Vergleichsvariablen des PFT (DUHM & HANSEN 1957) eingeschränkt Beziehungen aufweist. Erst bei Kontrolle intropunitiver und impunitiver

Reaktionstendenzen im PFT treten signifikante Beziehungen zwischen Extrapunitivität im PFT und den Aggressionsmotiv-Variablen des AM-GITTERs zutage. Es bestehen hier jedoch Möglichkeiten, die Zusammenhangsanalysen unter differenzierteren Gesichtspunkten zu komplettieren, sodaß sich das bislang bestehende Bild auch noch ändern könnte.

Als günstiger waren die Ausgangsbedingungen zum Nachweis divergenter Validität der Meßwerte des AM-GITTERs einzuschätzen. Wir haben jedoch versucht, uns diesen Nachweis soweit als möglich zu erschweren. Deswegen haben wir als Vergleichstest ein Verfahren ausgewählt, das von seiner konzeptuellen und theoretischen Gestaltungslogik her maximale Ähnlichkeit aufweist, jedoch einen anderen Motivbereich (Leistungsmotiv) zum Gegenstand hat, das LM-GITTER (SCHMALT 1976 b). Die Ergebnisse sprechen deutlich für divergente Validität der Meßwerte des AM-GITTERs; es bestehen keine übergreifenden Zusammenhänge zwischen den Meßwerten beider Verfahren. Diese Ergebnisse sind insbesondere deshalb als günstig zu betrachten, weil Zusammenhänge zwar theoretisch nicht gefordert, aber auch nicht gänzlich ausgeschlossen werden konnten. Es bestehen allerdings an einer ganz bestimmten Nahtstelle signifikante Zusammenhänge zwischen beiden GITTERn. Das LMG-Konstrukt FM 1 (Konzept mangelnder eigener Fähigkeit) ist signifikant mit unseren Meßwerten für Aggressions-Hemmung korreliert. Dieser Befund spricht jedoch nicht gegen, sondern eher für die Validität unserer Hemmungswerte. Der o.g. Zusammenhang könnte einerseits auf inhaltlichen Kongruenzen zwischen FM 1

und unserem Konstrukt "extrinsische Aggressions-Hemmung" beruhen, obgleich wir das nicht nachweisen konnten. Andererseits könnte darin lediglich zum Ausdruck kommen, daß verschiedene "Ängste" möglicherweise in unserer Gesellschaft einer gleichgerichteten Entwicklung unterliegen, die für die Kovariation verantwortlich zeichnet. Für letzteres gibt es empirische Hinweise (SCHMALT 1976 a, S. 127).

KORRELATION MIT AUSSENKRITERIEN: Nimmt man das Urteil Gleichaltriger zum Kriterium für aggressives und aggressionsgehemmtes Verhalten, so zeigen sich keine Zusammenhänge zwischen Außenkriterien und Meßwerten des AM-GITTERs. Oberflächlich betrachtet würde dieser Befund deutlich gegen die Validität unseres Verfahrens sprechen. Detailliertere Überlegungen zeigen jedoch, daß diese Schlußfolgerung keineswegs zwingend ist. Zunächst scheinen Verbesserungen in unserem Instrumentarium vonnöten, mit dem wir die Außenkriterien erhoben haben. Deren Ziel sollte sein, diese selbst valider zu erfassen. Sodann besteht der Verdacht, daß wir eine theoretisch relevante Moderatorvariable (Einstellungen und Praktiken des Elternhauses hinsichtlich der Aggression des Kindes) anzunehmen haben, d.h. mit differentieller Validität unseres Verfahrens rechnen müssen. Dies beinhaltet die Vorstellung, daß sich in den aufgefundenen Nullkorrelationen u.U. zwei einander entgegengerichtete Zusammenhänge überlagern. Empirische Hinweise, die diesen Verdacht stützen, finden wir bei LESSER (1957, 1959 a). Einer weiteren Überlegung zufolge waren diese

Ergebnisse aufgrund datentypologischer Inkongruenzen der korrelierten Variablen zu erwarten. Diese, einer Theorie von McCLELLAND (1951, 1980, 1981) entlehnte Vorstellung würde zwar bei Bestätigung des Verdachts auf differentielle Validität entkräftet, gab uns jedoch im vorliegenden Fall mit einen Hinweis darauf, die Außenkriterien unter einem neuen Aspekt zu betrachten, der datentypologische Kongruenz beinhaltete. Demzufolge untersuchten wir nachträglich die Außenkriterien auf Ebene des Urteilerverhaltens selbst. Kriterium war nun, in welcher Art und Weise ein Urteiler die abgefragten Verhaltensmerkmale zuschrieb (attribuierte). Bei dieser ex post facto - Analyse fanden wir nun deutliche Hinweise auf motivbedingte Attributionseffekte, die u.E. für die Konstruktvalidität der Meßwerte des AM-GITTERs sprechen. Jungen mit hohem Aggressionsmotiv neigen zu einer Über-, Jungen mit niedrigem Aggressionsmotiv zu einer Unterschätzung feindseliger Handlungsabsichten bei Gleichaltrigen, die ihnen bekannt sind. Bei Jungen mit hoher bzw. niedriger Aggressions-Hemmung dreht sich diese Richtungssystematik im Attributionsfehler um. Aufgrund empirischer Befunde (zusammenfassend PARKE & SLABY 1983 sowie DENGERINK 1976) ist die Annahme gerechtfertigt, daß diese Attributionstendenzen auch handlungswirksam sind.

GEMEINSAME FAKTORISIERUNG MIT AUSSENKRITERIEN: Die in Ergänzung zu der einfachen Korrelationsanalyse durchgeführte gemeinsame Faktorisierung der Meßwerte des AM-GITTERs mit Außenkriterien hat eine besondere Art der Übereinstimmungsvalidität nicht deutlich werden

lassen. Auf die Gründe dafür sind wir bereits eingegangen. Dagegen hat sich bei dieser Analyse gezeigt, daß wir unsere konstruktiven Zielsetzungen hinsichtlich der Dimensionalität der Meßwerte des AM-GITTERs als erreicht betrachten können.

Zusammenfassend läßt sich folgendes resümieren: Unsere ersten Untersuchungen zur Validität der Meßwerte des AM-GITTERs belegen positiv deren psychometrische Güteeigenschaften. Insofern kann man eine unabdingbare Voraussetzung für Validität als gegeben betrachten. Die Frage "Erfassen wir nun wirklich Merkmale des überdauernden Aggressionsmotivs und der überdauernden Aggressions-Hemmung von Kindern?" läßt sich vorläufig nur dahingehend beantworten, daß wir über erste Hinweise darauf verfügen, daß es so ist. Dies ersetzt nicht die Notwendigkeit, nach weiteren Hinweisen zu suchen. Ob der korrelationsstatistische Ansatz, dessen Möglichkeiten wir allerdings noch nicht voll ausgeschöpft haben, dazu den richtigen Rahmen bietet, ist nach den nunmehr vorliegenden Befunden begründetermaßen fraglich. Insgesamt weisen sie u.E. recht deutlich darauf hin, in einem neuen erneut nach Hinweisen zu suchen. Theoriegeleitet gestaltete Experimente mit Kindern, in die auf Personseite individuelle Differenzen im Aggressionsmotiv und in der Aggressions-Hemmung einbezogen sind, könnten aller Voraussicht nach mit größerer Aussicht auf Erfolg einen Beitrag zur Konstruktvalidität des AM-GITTERs erbringen, indem sie deren Handlungswirksamkeit in lebensnahen Situationen untersuchen (varianzanalytischer Ansatz). Dies bleibt jedoch nachfolgenden Untersuchungen überlassen.

4. Gesamtzusammenfassung

Ausgehend von KLEINADT's Motivationstheorie der Aggression wurde der Versuch unternommen, individuelle Differenzen im Aggressionsmotiv und in der Aggressions-Hemmung bei Kindern meßbar zu machen. Für die Entwicklung eines dazu geeigneten, testpsychologischen Verfahrens benutzten wir als methodischen Ansatz die GITTER-Technik (SCHMALT). Das Testkonzept der GITTER-Technik sieht zu einem Satz motivthematisch relevanter Bildsituationen zur Darstellung situativer Aspekte motivierten Handelns eine Reihe motivthematisch bedeutsamer Aussagen vor, mit denen die dabei beteiligten personseitigen Aspekte zum Ausdruck gebracht werden sollen und denen demzufolge der Status motivationaler Konstrukte beigemessen wird. Die Aufgabe des Probanden besteht darin, durch Ankreuzen der für ihn in einer bestimmten Bildsituation bedeutsamen Motivkonstruktindikatoren (Aussagen) eine für ihn typische, motivgerichtete Auffassung und Deutung dieser Situation zum Ausdruck zu bringen und damit die Grundlage für einen Einblick in sein individuelles Motivstrukturgefüge zu legen.

Vor allem aufgrund der Erprobungscharakters der vorliegenden Arbeit beschränkten wir uns auf die Entwicklung eines Verfahrens, das für Jungen im Alter von etwa 9 bis 11 Jahren Geltung haben sollte. Zur Darstellung eines für Aggression bedeutsamen Ausschnitts aus dem sozialen Umfeld von Kindern in einem "Aggressionsmotiv-Gitter" (AM-Gitter) wählten wir die Bildsituationen eines Aggressions-TAT für Jungen (LESSER), die bereits für

Aggression zwischen Gleichaltrigen skaliert und geprüft worden waren. Die Bilder wurden dann erstmals von uns an deutschsprachigen Kindern auf ihre grundsätzliche Verwendbarkeit hin untersucht und am Ende 6 davon - darunter ein modifiziertes - als situationsseitiger Bestandteil des AM-Gitters beibehalten.
Die personseitigen Bestandteile des AM-Gitters entwickelten wir nach inhaltsanalytischen Vorgaben eines Aggressions-TAT für Erwachsene (KORNADT) als eine Art von Musterlösungen für die Anwendung aggressionsmotiv- und aggressionshemmungsspezifischer Kategorien in der Form kleiner, ganzer Sätze, die auf die o.g. Bildsituationen inhaltlich zu beziehen waren. Diese Aussagen wurden zunächst nach Expertenurteil und anschließend nach psychometrischen Kriterien weiter geprüft und am Ende eine Auswahl von 26 Aussagen (jeweils 13 für eine Motivkomponente) beibehalten.

Für die Auswertung des auf dieser Grundlage erstellten AM-Gitters wurden insgesamt 8 Subskalen vorgeschlagen. Zur Standardauswertung sind das im einzelnen die Skalen A (Aggressionsmotiv), AH (Aggressions-Hemmung), n-A. (Netto-Aggression) und Ges.-A. (Gesamt-Aggressivität). Zur differentiellen Auswertung wurden die Skalen ADM (bedürfnisbezogene Bestandteile des Aggressionsmotivs), ADH (handlungsbezogene Bestandteile des Aggressionsmotivs), AHI (intrinsische Aggressions-Hemmung) und AHE (extrinsische Aggressions-Hemmung) berechnet.

Die Meßwerte der Skalen des AM-Gitters sind im Rahmen eines Generalfaktormodells beschreibbar, das aus dem unipolaren Faktor AGGRESSION

und dem bipolaren Richtungsfaktor SOZIALE WERTUNG aufgebaut ist. Den 8 Subskalen entsprechen unterschiedlich gerichtete Subtestvektoren innerhalb des o.g. Modells. Die Position der Subtestvektoren der Standardskalen des AM-Gitters konnte dabei theoretisch festgelegt werden.

Prüfungen der Meßgenauigkeit der Skalen nach verschiedenen Ansätzen erbrachten zufriedenstellende Ergebnisse. Zusammenfassend zu den Entwicklungsarbeiten vgl. die Abschnitte 1.4 und 2.9 des empirischen Teils dieser Arbeit.

Erste Untersuchungen zur Prüfung der Validität des AM-Gitters lassen den Schluß zu, daß wir mit unseren Skalen überdauernde Aspekte des Aggressionsmotivs und der Aggressions-Hemmung bei Kindern erfassen (zusammenfassend vgl. Abschnitt 3.6 des empirischen Teils dieser Arbeit). Auf der Ebene individueller sowie gruppenbezogener Kennwerte können wir bei einer Wiederholungsmessung nach 8 Wochen bestehende Zusammenhänge zufallskritisch absichern. Zu Testmessungen, die andere Persönlichkeitsmerkmale erfassen, weisen unsere Kennwerte im Grunde keine Beziehung auf. Eine Ausnahme bildet eine schwache Korrelation zwischen unseren Kennwerten für Aggressions-Hemmung und einem Kennwert des LM-Gitters für Kinder (SCHMALT), mit dem die Ausprägung des Konzepts mangelnder eigener Fähigkeit hinsichtlich leistungsthematischer Zielsetzungen (FM 1) beschrieben wird. Sollten bestimmte theoretische Überlegungen zutreffen, könnte dies jedoch gerade für die Validität unserer Kennwerte sprechen. Zu Testmessungen, die dem gleichen Gegenstandsbereich (kindliche Aggressivität)

angehören, zeigen sich nur sehr eingeschränkt Beziehungen. Diese waren bei der streng theoriegeleiteten Entwicklung unseres Verfahrens als auch bei dem relativ unklaren Gültigkeitsbereich der Vergleichstests nicht verwunderlich. Die Berechnungen zur "konvergenten" Validität sind demzufolge unter einem mehr pragmatischen Gesichtspunkt (Standortbestimmung) zu sehen. Sie zeigen, daß wir verschiedene Aspekte kindlicher Aggressivität erfassen, die offensichtlich einer anderen Ebene (McCLELLAND) angehören als diejenigen, die mit dem PICTURE-FRUSTRATION-TEST oder mit verschiedenen Subskalen des PFK 9-14 erfaßt werden.
Zu Außenkriterien für Aggression und Aggressions-Hemmung (Ratings) konnten vorerst noch keine Beziehungen aufgezeigt werden; das Einschätzverhalten bei der Abgabe der Außenkriterien unterliegt dagegen einer aggressionsmotiv- und aggressionshemmungsspezifischen Ausrichtung, die als motivbedingter Attributionseffekt aufgefaßt werden kann und für die Konstruktvalidität des AM-Gitters spricht.

5. Literatur

Angleitner, A.: Einführung in die Persönlichkeitspsychologie Bd. 1. Nichtfaktorielle Ansätze. Huber, Bern 1980

Angst, W.: Aggression bei Affen und Menschen. Springer, Berlin 1980

Auld, F., Eron, L.D. & Laffal, J.: Application of Guttman's scaling method to the TAT. Educational and Psychological Measurement, 1955, 15, 422-435

Bäumler, G. & Breitenbach, W.: Zusammenhänge zwischen Intelligenz, Konzentration, Angst und Leistungsmotivation bei einer studentischen Stichprobe. Psychologie und Praxis, 1970, 14, 37-40

Bäumler, G. & Dvorak, H.P.: Weitere Untersuchungen zum Zweifaktorenmodell der Leistungsmotivation. Psychologie und Praxis, 1969, 13, 122-138

Bäumler, G. & Weiss, R.: Eine Zweifaktorentheorie der nach der TAT-Methode gemessenen Leistungsmotivation (Heckhausen). Psychologie und Praxis, 1967, 11, 23-45

Bandura, A.: Social learning theory of aggression. In: Knutson, J.F. (Ed.): The control of aggression, 201-250. Aldine, Chicago 1973

Bandura, A. & Walters, R.: Social learning and personality development. Holt, Rinehart & Winston, New York 1963

Baron, R.: Human aggression. Plenum Press, New York 1977

Belschner, W., Lischke, G. & Selg, H.: Foto-Hand-Test (FHT) zur Erfassung der Aggressivität. Alber, Freiburg 1971

Bem, D. & Allen, A.: On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. Psychological Review, 1974, 81, 506-520

Berkowitz, L.: Aggression. McGraw-Hill, New York 1962

- Some determinants of impulsive aggression. Psychological Review, 1974, 81, 165-176

Beutel, P. & Schubö, W.: SPSS 9. Statistik-Programm-System für die Sozialwissenschaften. Fischer, Stuttgart 1983

Birbaumer, N.: Angst. In: Herrmann, Th. et al. (Hrsg.): Handbuch psychologischer Grundbegriffe, 27-38. Kösel, München 1977

Bowers, K.S.: Situationism in psychology: An analysis and a critique. Psychological Review, 1973, 80, 307-336

Brickenkamp, R. (Hrsg.): Handbuch psychologischer und pädagogischer Tests. Hogrefe, Göttingen 1975

- Erster Ergänzungsband zum Handbuch psychologischer und pädagogischer Tests. Hogrefe, Göttingen 1983

Bricklin, B., Piotrowski, Z.A. & Wagner, E.E.: The hand test. A new projective test with special reference to the prediction of overt aggressive behavior. Thomas, Springfield Ill. 1962

Burkhardt, K.: Zur Messung des Aggressionsmotivs von Kindern. Magister-Arbeit. FR Allgemeine Erziehungswissenschaft der Universität des Saarlandes, Saarbrücken 1978 (unveröffentlicht)

Buss, A.: The psychology of aggression. Wiley, New York 1961

Campbell, D.T. & Fiske, D.W.: Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105

Christoph-Lemke, Ch.: Aggression. Theorie und Diagnostik. Akademische Buchhandlung GmbH, München 1972

Clauß, G. & Ebner, H.: Grundlagen der Statistik. Deutsch, Zürich 1975

Cronbach, L.J.: Essentials of psychological testing. Harper & Row, New York 1970

Davison, G.C. & Neale, J.M.: Klinische Psychologie. Urban & Schwarzenberg, München 1974

Dengerink, H.A.: Personality variables as mediators of attack-instigated aggression. In: Geen, R.G. & O'Neal, E.C. (Eds.): Perspectives on aggression, 61-98. Academic Press, New York 1976

Dieterich, R.: Psychodiagnostik. Reinhardt, München 1973

Dodge, K.A.: Social cognition and children's aggressive behavior. Child Development, 1980, 51, 162-170

Dodge, K.A. & Frame, C.L.: Social cognitive biases and deficits in aggressive boys. Child Development, 1982, 53, 620-635

Dollard, J., Doob, L.W., Miller, N.E., Mowrer, O.H. & Sears, R. R.: Frustration and aggression. Yale University Press, New Haven Conn. 1939

Duhm, E. & Hansen, J.: Der Rosenzweig P-F-Test (Form für Kinder). Hogrefe, Göttingen 1957

Elizur, A.: Content analysis of the Rorschach with regard to anxiety and hostility. Journal of Projective Techniques and Personality Assessment, 1949, 13, 247-284

Eron, L.D., Walder, L.O., Huesmann, L.R. & Lefkowitz, M.M.: The convergence of laboratory and field studies of the development of aggression. In: De Wit, J. & Hartup, W.W. (Eds.): Determinants and origins of aggressive behavior, 347-380. Mouton, The Hague 1974

Feshbach, S.: Aggression. In: Mussen, P.H. (Ed.): Carmichael's manual of child psychology, Vol. 2, 159-259. Wiley, New York 1970

Fisch, R.: Achievement Motivation Matrix. An instrument for the assessment of the intensity and extensity of achievement motivation. Paper presented at the 2nd Prague Conference on Psychology of Human Learning and Problem Solving, July 16-20, Prague, Czechoslovakia 1973

- Anleitung zum Gebrauch der LM-Matrix. Ein Verfahren zur Erfassung der Intensität und Extensität der Leistungsmotivation bei Jugendlichen und Erwachsenen. Arbeiten der Fachrichtung Psychologie, Universität des Saarlandes, Saarbrücken 1974

Fisch, R., Schmalt, H.D. & Fisch, H.: Bericht über die Entwicklung eines Meßverfahrens zur Bestimmung der Intensität und Extensität der Leistungsmotivation. Unveröffentlichter Arbeitsbericht, Psychologisches Institut der Universität des Saarlandes und Psychologisches Institut der Ruhr-Universität Bochum, Saarbrücken und Bochum 1970

Fromm, E.: Anatomie der menschlichen Destruktivität. Deutsche Verlags-Anstalt, Stuttgart 1974

Fruchter, B.: Introduction to factor analysis. Van Nostrand, Princeton New Jersey 1954

Füllgrabe, U.: Kriminalpsychologie. Verlag für Angewandte Psychologie, Stuttgart 1983

Gaensslen, H. & Schubö, W.: Einfache und komplexe statistische Analyse. Reinhardt, München 1973

Geen, R.G. & O'Neal, E.L. (Eds.): Perspectives on aggression. Academic Press, New York 1976

GMD: Programm-Katalog Nr. 1 RZ-Darmstadt. Gesellschaft für Mathematik und Datenverarbeitung mbH. Bonn, Rechenzentrum Darmstadt, Darmstadt 1976

Goldfried, M.R. & Kent, R.N.: Traditional versus behavioral personality assessment: a comparison of methodological and theoretical assumptions. Psychological Bulletin, 1972, 77, 409-420

Guilford, J.P.: Psychometric methods. McGraw-Hill, New York 1954

Halisch, F.: Über die Güteeigenschaften des LM-Gitters für Kinder (Schmalt). Kritische Anmerkungen zu einer Analyse von Krug & Walter (1979). Diagnostica, 1982, 28, 146-153

Heckhausen, H.: Hoffnung und Furcht in der Leistungsmotivation. Hain, Meisenheim am Glan 1963

- Vorbemerkungen des Herausgebers. In: Zumkley, H.: Aggression und Katharsis, 5-6. Hogrefe, Göttingen 1978

- Motivation und Handeln. Springer, Berlin 1980

Heinrich, K.: Filmerleben - Filmwirkung - Filmerziehung. Schroedel, Hannover 1961

Hilke, R. & Kempf, W. (Hrsg.): Aggression. Naturwissenschaftliche und kulturwissenschaftliche Perspektiven der Aggressionsforschung. Huber, Bern 1982

Hiltmann, H. & Luhr, R.: Wortassoziation und verbale Ergänzungsverfahren. In: Groffmann, K.J. & Michel, L. (Hrsg.): Enzyklopädie der Psychologie. Psychologische Diagnostik Bd. 3, Persönlichkeitsdiagnostik, 373-429. Hogrefe, Göttingen 1982

Jüttemann, G.: "Aggression" als wissenschaftlicher Begriff: Versuch einer Explikation. In: Hilke, R. & Kempf, W. (Hrsg.): Aggression. Naturwissenschaftliche und kulturwissenschaftliche Perspektiven der Aggressionsforschung, 281-316. Huber, Bern 1982

Kastner, M.: Einige Erlebens- und Verhaltensvariablen als Determinanten der Berufsentscheidung. In: Feger, H. (Hrsg.): Studien zur intraindividuellen Konfliktforschung, 205-298. Huber, Bern 1977

Kenny, D. & Bijou, S.: Ambiguity of pictures and extent of personality factors in fantasy responses. Journal of Consulting Psychology, 1953, 17, 283-288

Koeck, R. & Strube, G.: Beobachtung und Befragung. In: Strube, G. (Hrsg.): Die Psychologie des 20. Jahrhunderts Band V. Binet und die Folgen, 151-212. Kindler, Zürich 1977

Kornadt, H.J.: Toward a motivation theory of aggression and aggression inhibition: Some considerations about an aggression motive and their application to TAT and catharsis. In: De Wit, J. & Hartup, W.W. (Eds.): Determinants and origins of aggressive behavior, 567-577. Mouton, The Hague 1974

- Grundzüge einer Motivationstheorie der Aggression. Arbeitsberichte aus der FR Allgemeine Erziehungswissenschaft Nr. 5. Universität des Saarlandes, Saarbrücken 1980

- Die Entwicklung der Frustrations- und der Aggressionsforschung. In: Kornadt, H.J. (Hrsg.): Aggression und Frustration als psychologisches Problem, Bd. 1, 3-59. Wissenschaftliche Buchgesellschaft, Darmstadt 1981 (a)

- Aggression und Frustration als psychologisches Problem, Bd. 1. Wissenschaftliche Buchgesellschaft, Darmstadt 1981 (b)

- Aggressionsmotiv und Aggressionshemmung, Bd. 1 (a), Bd. 2 (b). Huber, Bern 1982

  Grundzüge einer Motivationstheorie der Aggression. In: Hilke, R. & Kempf, W. (Hrsg.): Aggression. Naturwissenschaftliche und kulturwissenschaftliche Perspektiven der Aggressionsforschung, 86-111. Huber, Bern 1982 (c)

- Motivation theory of aggression and its relation to social psychological approaches. In: Mummendey, A. (Ed.): Social psychology of aggression. From individual behavior to social interaction, 21-31. Springer, Berlin 1984

Kornadt, H.J., Eckensberger, L.H. & Emminghaus, W.B.: Cross-cultural research on motivation and its contribution to a general theory of motivation. In: Triandis, H.C. & Lonner, W. (Eds.): Handbook of cross-cultural psychology, Vol. 3, Basic processes, 223-321. Allyn & Bacon, Boston 1980

Kornadt, H.J. & Zumkley, H.: Thematische Apperzeptionsverfahren. In: Groffmann, K.J. & Michel, L. (Hrsg.): Enzyklopädie der Psychologie. Psychologische Diagnostik Bd. 3, Persönlichkeitsdiagnostik, 258-272. Hogrefe, Göttingen 1982

Kornadt, H.J., Zumkley, H. & Burkhardt, K.: Das Aggressions-Motiv-Gitter. Ein objektives Verfahren zur Messung des Aggressions-Motivs und der Aggressions-Hemmung bei Kindern: Konstruktion und erste Ergebnisse. Arbeitsberichte aus der FR Allgemeine Erziehungswissenschaft Nr. 12. Universität des Saarlandes, Saarbrücken 1982

Krug, S.: Förderung und Änderung des Leistungsmotivs: Theoretische Grundlagen und deren Anwendung. In: Schmalt, H.D. & Meyer, W.U. (Hrsg.): Leistungsmotivation und Verhalten, 221-247. Klett, Stuttgart 1976

Krug, S. & Hanel, J.: Motivänderung: Erprobung eines theoriegeleiteten Trainingsprogrammes. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 1976, 8, 274-287

Kuhl, J.: Situations-, reaktions- und personbezogene Konsistenz des Leistungsmotivs bei der Messung mittels des Heckhausen-TAT. Archiv für Psychologie, 1978, 130, 37-52

Lazarus, R.S.: Ambiguity and non-ambiguity in projective testing. Journal of Abnormal and Social Psychology, 1953, 48, 443-445

Lefkowitz, L., Eron, L., Walder, L. & Huesmann, L.: Growing up to be violent: a longitudinal study of the development of aggression. Pergamon, New York 1977

Lesser, G.S.: The relationship between overt and fantasy aggression as a function of maternal response to aggression. Journal of Abnormal and Social Psychology, 1957, 55, 218-221

- Conflict analysis of fantasy aggression. Journal of Personality, 1958, 26, 29-41 (a)

- Application of Guttman's scaling method to aggressive fantasy in children. Educational and Psychological Measurement, 1958, 18, 543-551 (b)

- Population differences in construct validity. Journal of Consulting Psychology, 1959, 23, 60-65 (a)

- The relationship between various forms of aggression and popularity among lower class children. Journal of Educational Psychology, 1959, 50, 20-25 (b)

Leyhausen, P.: Zur Naturgeschichte der Angst (1967). In: Lorenz, K. & Leyhausen, P.: Antriebe tierischen und menschlichen Verhaltens. Gesammelte Abhandlungen, 272-296. Piper, München 1973

Lienert, G.: Testaufbau und Testanalyse. Beltz, Weinheim 1969

Lorenz, K.: Das sogenannte Böse. Zur Naturgeschichte der Aggression. Borotha-Schoeler, Wien 1963

Maccoby, E.E.: Social development. Harcourt Brace Jovanovich, New York 1980

MacLean, P.D.: The limbic system with respect to self-preservation and the preservation of the species. Journal of Nervous and Mental Disease, 1958, 127, 1-11

Magnusson, D. & Endler, N.S.: Interactional psychology: present status and future prospects. Reports from the department of psychology, University of Stockholm, Stockholm 1976

McClelland, D.C.: Personality. Holt, Rinehart & Winston, New York 1951

- Methods of measuring human motivation. In: Atkinson, J.W. (Ed.): Motives in fantasy, action, and society, 7-42. Van Nostrand, New York 1958

- Motive dispositions. In: Wheeler, L. (Ed.): Review of personality and social psychology Vol. 1, 10-41. Sage Publications, Beverly Hills 1980

- Is personality consistent? In: Rabin, A., Aronoff, J., Barclay, A. & Zucker, R. (Eds.): Further explorations in personality, 87-113. Wiley, New York 1981

Mees, U.: Modifikation "aggressiven" Verhaltens. In: Selg, H.: Menschliche Aggressivität, 147-172. Hogrefe, Göttingen 1974

- Wie können Aggressionen im Alltag empirisch ermittelt werden? In: Hilke, R. & Kempf, W. (Hrsg.): Aggression. Naturwissenschaftliche und kulturwissenschaftliche Perspektiven der Aggressionsforschung, 355-360. Huber, Bern 1982

Mees, U. & Fieguth, G.: Sequentielle Beobachtung und Analyse aggressiven Kindverhaltens. In: Tack, W.H. (Hrsg.): Bericht über den 30. Kongreß der Deutschen Gesellschaft für Psychologie in Regensburg 1976, Bd. 2, 77-79. Hogrefe, Göttingen 1977

Michaelis, W.: Verhalten ohne Aggression? Versuch zur Integration der Theorien. Kiepenheuer & Witsch, Köln 1976

Miller, N.E.: The frustration-aggression hypothesis. Psychological Review, 1941, 48, 337-342

Mischel, W.: Personality and assessment. Wiley, New York 1968

- Toward a cognitive social learning reconceptualization of personality. Psychological Review, 1973, 80, 252-283

Mitscherlich, A.: Die Idee des Friedens und die menschliche Aggressivität. Vier Versuche. Suhrkamp, Frankfurt 1969

Moore, M.: Aggression themes in a binocular rivalry situation. Journal of Personality and Social Psychology, 1966, 3, 685-688

Moosbrugger, H.: Multivariate statistische Analyseverfahren. Kohlhammer, Stuttgart 1978

Moyer, K.E.: The physiology of hostility. Markham, Chicago 1971 (a)

- Experimentelle Grundlagen eines physiologischen Modells aggressiven Verhaltens. In: Schmidt-Mummendey, A. & Schmidt, H.D. (Hrsg.): Aggressives Verhalten. Neue Ergebnisse der psychologischen Forschung, 25-57. Juventa, München 1971 (b)

- The physiological inhibition of hostile behavior. In: Knutson, J.F. (Ed.): The control of aggression, 9-38. Aldine, Chicago 1973

Mummendey, A.: Zum Nutzen des Aggressionsbegriffs für die psychologische Aggressionsforschung. In: Hilke, R. & Kempf, W. (Hrsg.): Aggression. Naturwissenschaftliche und kulturwissenschaftliche Perspektiven der Aggressionsforschung, 317-333. Huber, Bern 1982

Mummendey, A., Linneweber, V. & Löschper, G.: Aggression: from act to interaction. In: Mummendey, A. (Ed.): Social psychology of aggression. From individual behavior to social interaction, 69-106. Springer, Berlin 1984

Murray, H.A.: Thematic apperception test manual. Harvard University Press, Cambridge 1943

Mussen, P.H., Conger, J.J. & Kagan, J.: Lehrbuch der Kinderpsychologie. Klett-Cotta, Stuttgart 1981

Mussen, P.H. & Naylor, H.K.: The relationships between overt and fantasy aggression. Journal of Abnormal and Social Psychology, 1954, 49, 235-240

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K. & Bent, D.H.: Statistical package for the social sciences. McGraw-Hill, New York 1975

Olweus, D.: Prediction of aggression. Scandinavian Test Corporation, Stockholm 1969

- Personality factors and aggression: With special reference to violence within the peer group. In: De Wit, J. & Hartup, W.W. (Eds.): Determinants and origins of aggressive behavior, 535-565. Mouton, The Hague 1974

- Aggression in the schools: Bullies and whipping boys. Hemisphere, Washington D.C. 1978

- Stability of aggressive reaction patterns in males: A review. Psychological Bulletin, 1979, 86, 852-875

- Familial and temperamental determinants of aggressive behavior in adolescent boys: A causal analysis. Developmental Psychology, 1980, 16, 644-660

Parens, H.: The development of aggression in early childhood. Jason Aronson, New York 1979

Parke, R.D. & Slaby, R.G.: The development of aggression. In: Mussen, P.H. (Ed.): Child Psychology Vol. 4. Socialization, Personality and Social Development, 547-641. Wiley, New York 1983

Patterson, G.R.: Treatment for children with conduct problems: a review of outcome studies. In: Feshbach, S. & Fraczek, A. (Eds.): Aggression and behavior change. Biological and sociological processes, 83-132. Praeger, New York 1979

Petermann, F. & Petermann, U.: Training mit aggressiven Kindern. Urban & Schwarzenberg, München 1978

- Erfassungsbogen für aggressives Verhalten in konkreten Situationen. Westermann, Braunschweig 1980

Purcell, K.: The TAT and and antisocial behavior. Journal of Consulting Psychology, 1956, 20, 449-456

Rauchfleisch, U.: Handbuch zum Rosenzweig Picture-Frustration -Test (PFT). 2 Bände. Huber, Bern 1979

- Testpsychologie. Vandenhoek & Ruprecht, Göttingen 1980

Rey, E.R.: Allgemeine Probleme psychologischer Tests. In: Strube, G. (Hrsg.): Die Psychologie des 20. Jahrhunderts Band V. Binet und die Folgen, 65-101. Kindler, Zürich 1977

Rösler, F., Manzey, D., Jesse, J. & Grau, U.: Geschlechts- und formulierungsbedingte Unterschiede beim LM-Gitter für Kinder (Schmalt). Diagnostica, 1980, 26, 244-250

- Ist das LM-Gitter nur ein LM-Test? Eine dreimodale Faktorenanalyse des LM-Gitters für Kinder (Schmalt). Diagnostica, 1982, 28, 131-145

Rohkamm, R.: Zur Dimensionalität der Aggressivität. Fachbereich Medizin der Universität Hamburg. Hamburg 1974 (unveröffentlicht)

Rollett, B. & Bartram, M.: Anstrengungsvermeidungstest. Westermann, Braunschweig 1977

Rorschach, H.: Psychodiagnostik. Methodik und Ergebnisse eines wahrnehmungsdiagnostischen Experiments (Deutenlassen von Zufallsformen). Huber, Bern 1962 (Erstausgabe 1921)

Sachs, L.: Statistische Auswertungsmethoden. Springer, Berlin 1972

Schmalt, H.D.: Die Gitter-Technik - ein objektives Verfahren zur Messung des Leistungsmotivs bei Kindern. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 1973, 5, 231-252

- Entwicklung und Validierung einer neuen Technik zur Messung verschiedener Aspekte des Leistungsmotivs - das LM-GITTER. Dissertation, Philosophische Fakultät der Ruhr-Universität Bochum, Bochum 1974 (unveröffentlicht)

- Die Messung des Leistungsmotivs. Hogrefe, Göttingen 1976 (a)

- Das LM-Gitter. Hogrefe, Göttingen 1976 (b)

- Machtmotivation. Psychologische Rundschau, 1979, 30, 269-285

Schmidtchen, S.: Psychologische Tests für Kinder und Jugendliche. Hogrefe, Göttingen 1975

Sears, R.R.: Personality development in the family. In: Seidman, J.M. (Ed.): The child, 117-137. Holt, Rinehart & Winston, New York 1958

Seitz, W. & Rausche, A.: Persönlichkeitsfragebogen für Kinder 9-14. Westermann, Braunschweig 1976

Süllwold, F.: Die Entwicklung von Attitüden-Skalen zur Messung von Aggressionstendenzen bei Kindern und Jugendlichen. Schweizerische Zeitschrift für Psychologie, 1959, 18, 179-188

- Konstruktion, Verläßlichkeitsbestimmung und Validation von Attitüden-Skalen zur Erfassung der Aggressionstendenzen bei Jugendlichen. In: Heinrich, K.: Filmerleben - Filmwirkung - Filmerziehung, 111-167. Schroedel, Hannover 1961

Selg, H.: Menschliche Aggressivität. Hogrefe, Göttingen 1974

Steinhausen, H.: Untersuchungen zur Validität des Rosenzweig-Picture-Frustration-Test (RPFT) bei Kindern und Jugendlichen. Diagnostica, 1979, 25, 248-256

Straus, M.A.: A general systems theory approach to the development of a theory of violence between family members. In: De Wit, J. & Hartup, W.W. (Eds.): Determinants and origins of aggressive behavior, 401-412. Mouton, The Hague 1974

Tack, W.H.: Scriptum zur Übung Testtheorie. Fachrichtung Psychologie der Universität des Saarlandes, Saarbrücken 1973 (unveröffentlicht)

Thurstone, L.L.: Multiple-factor analysis. The University of Chicago Press, Chicago 1969

Toch, H.H. & Schulte, W.H.: Readiness to perceive violence as a result of police training. British Journal of Psychology, 1961, 52, 389-393

Torgerson, W.S.: Theory and methods of scaling. Wiley, New York 1958

Überla, K.: Faktorenanalyse. Springer, Berlin 1977

Wachtel, P.: Psychoanalysis and behavior therapy: Toward an integration. Basic Books, New York 1977

Wagner, H.: Hamburger Verhaltensbeurteilungsliste (HAVEL). Hogrefe, Göttingen 1981

Wagner, H. & Baumgärtel, F.: Hamburger Persönlichkeitsfragebogen für Kinder (HAPEF-K). Hogrefe, Göttingen 1978

Weizsäcker, C.F. v.: Die Einheit der Natur. Deutscher Taschenbuch Verlag, München 1982

Westermann, R. & Hager, W.: Interindividuelle Unterschiede bei der Beurteilung von Situationen im Rosenzweig-P-F-Test. Diagnostica, 1983, 24, 220-235

Westhoff, K.: Beziehungen zwischen Neurotizismus und Merkmalen intraindividueller Konflikte. In: Feger, H. (Hrsg.): Studien zur intraindividuellen Konfliktforschung, 149-204. Huber, Bern 1977

Wieczerkowski, W., Nickel, H., Janowski, A., Fittkau, B. & Rauer, W.: Angstfragebogen für Schüler. Westermann, Braunschweig 1975

Wilson, J.P.: Motives, values and moral judgements. Journal of Personality Assessment, 1983, 47, 414-426

Windelband, W. & Heimsoeth, H.: Lehrbuch der Geschichte der Philosophie. Mohr, Tübingen 1957

Zegans, L.S.: Towards a unified theory of human aggression. British Journal of Medical Psychology, 1971, 44, 355-365

Zillmann, D.: Hostility and aggression. Erlbaum, Hillsdale New Jersey 1979

Zumkley, H.: Aggression und Katharsis. Hogrefe, Göttingen 1978

- Zur Gültigkeit von Banduras modeling-Paradigma für die Aggressionsforschung. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 1980, 12, 30-39

- Aggressionskontrolle - Implikationen lerntheoretischer Konzepte aggressiven Verhaltens. Zeitschrift für personenzentrierte Psychologie und Psychotherapie, 1983, 2, 153-169

- Einflüsse von individuellen Differenzen der Aggressivität bei kausaler Ambiguität. Zeitschrift für Differentielle und Diagnostische Psychologie, 1984, 5, 131-140

- Individual differences and aggressive interactions. In: Mummendey, A. (Ed.): Social psychology of aggression. From individual behavior to social interaction, 33-49. Springer, Berlin 1984

- Operante versus respondente Messung von Aggressivität, Vigilanz und Wahrnehmungsabwehr bei der Identifikation von aggressionsrelevanten Wörtern. Psychologische Beiträge, 1985, Heft 1/2, in Vorbereitung.

6. Anhang

I: Zusammenstellung der zur Entwicklung des AM-GITTERs verwendeten Bildsituationen.

Nr. 1 bis Nr. 10: Bilder des LESSER-TAT in identischer Numerierung;

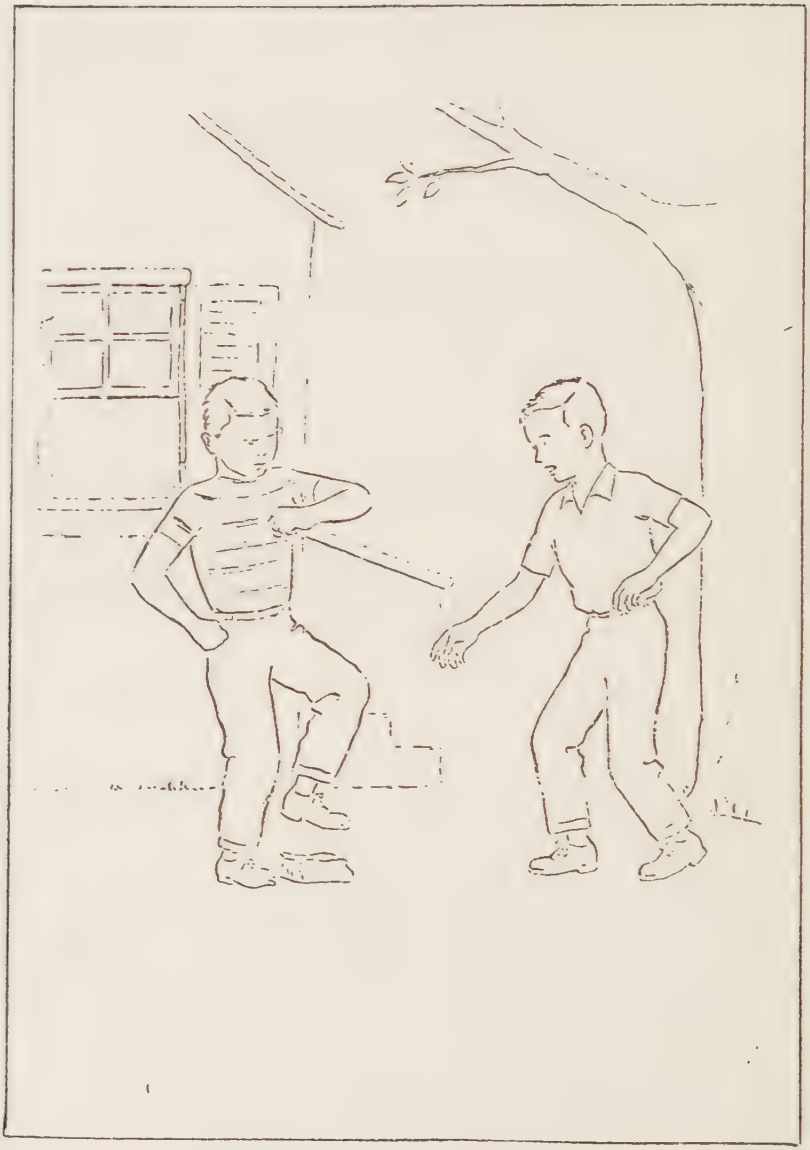
Nr. 11: eigene Modifikation des LESSER-Bildes Nr. 3

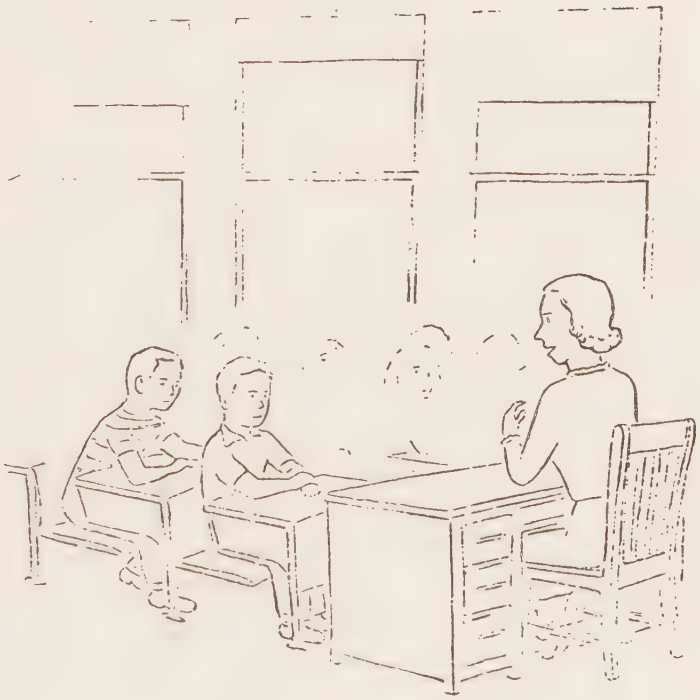
II: Liste der zur Entwicklung des AM-GITTERs verwendeten konstruktthematischen und neutralen Aussagen.

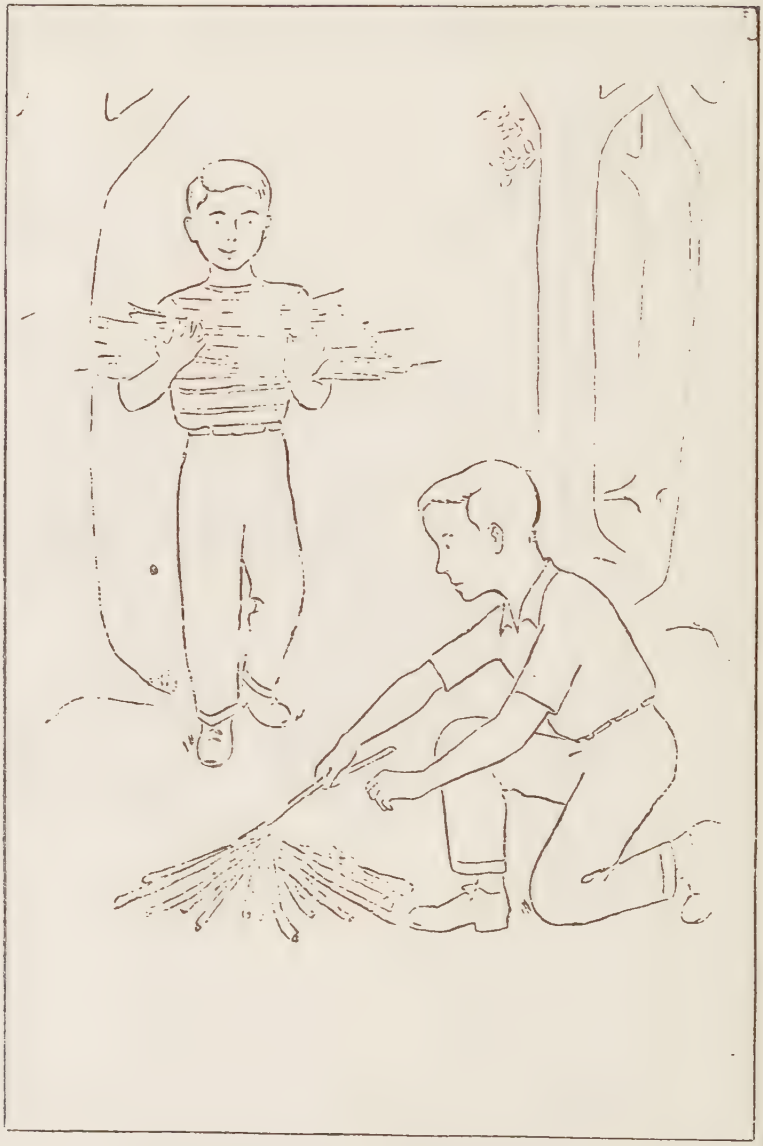
III: Anmerkungen zum Text.

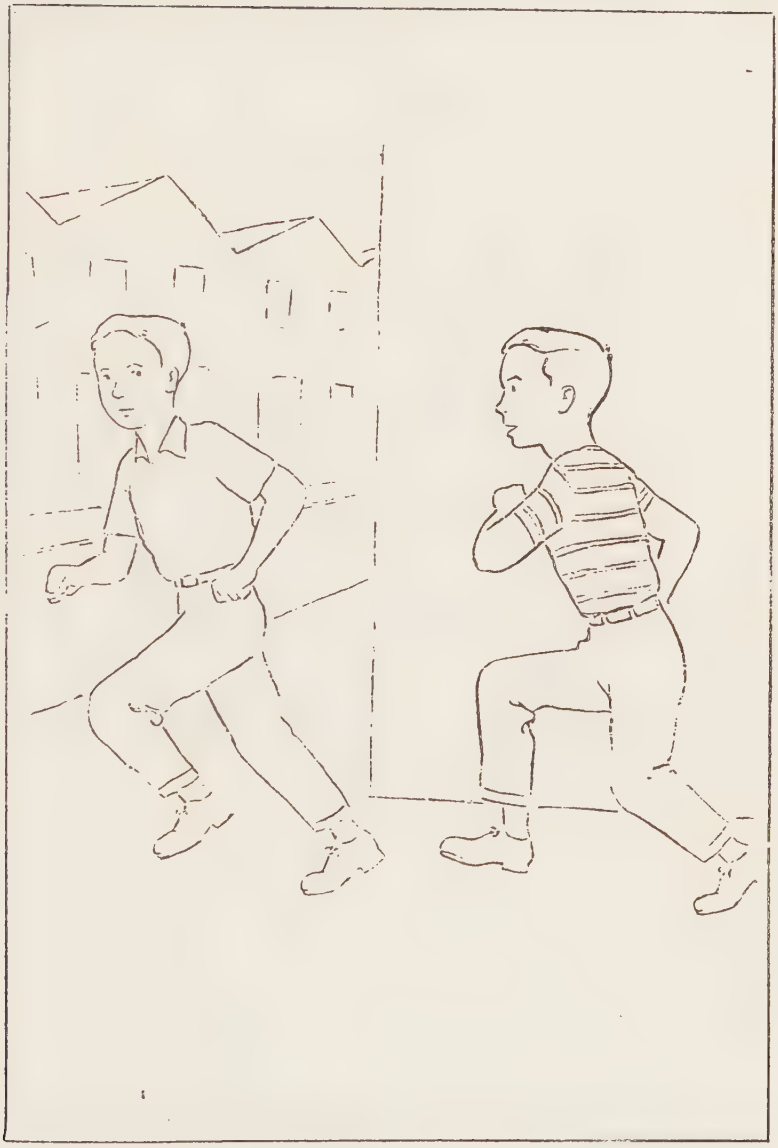
IV: Rohwerteverteilungen der Skalen des AM-GITTERs.

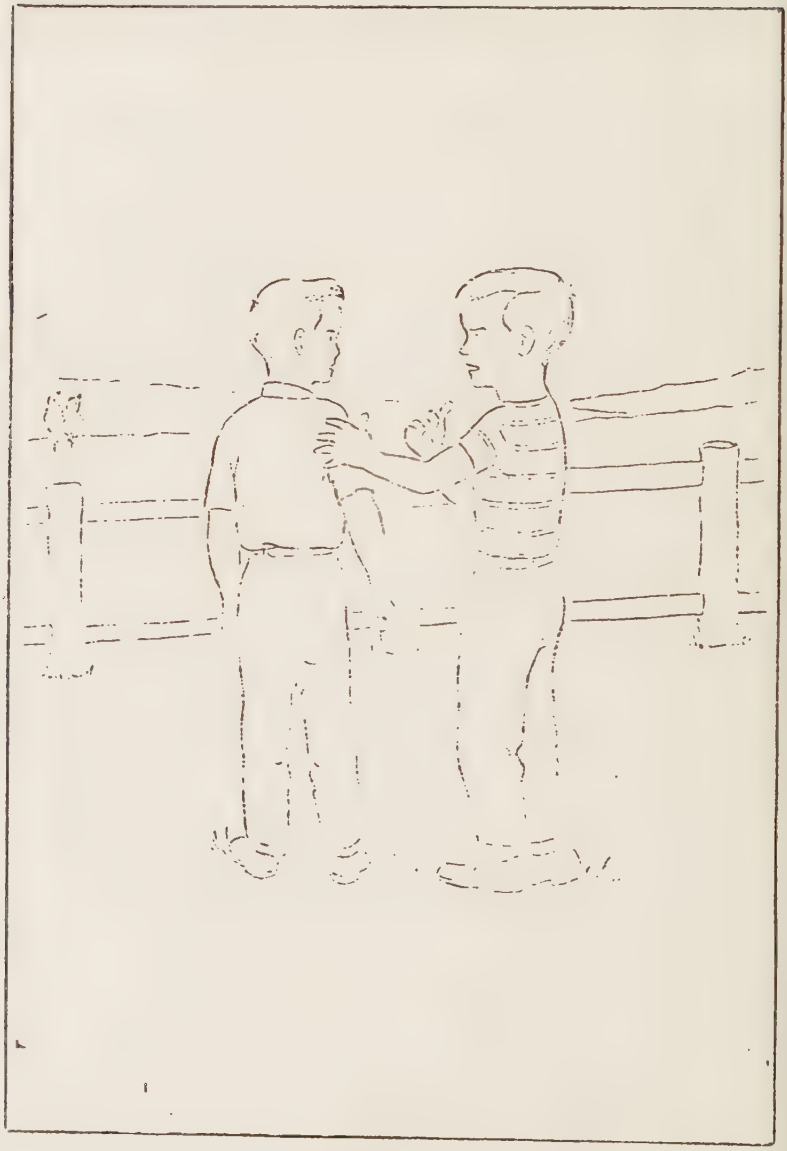
V: Instruktion und Skalen zum Gleichaltrigen-Rating

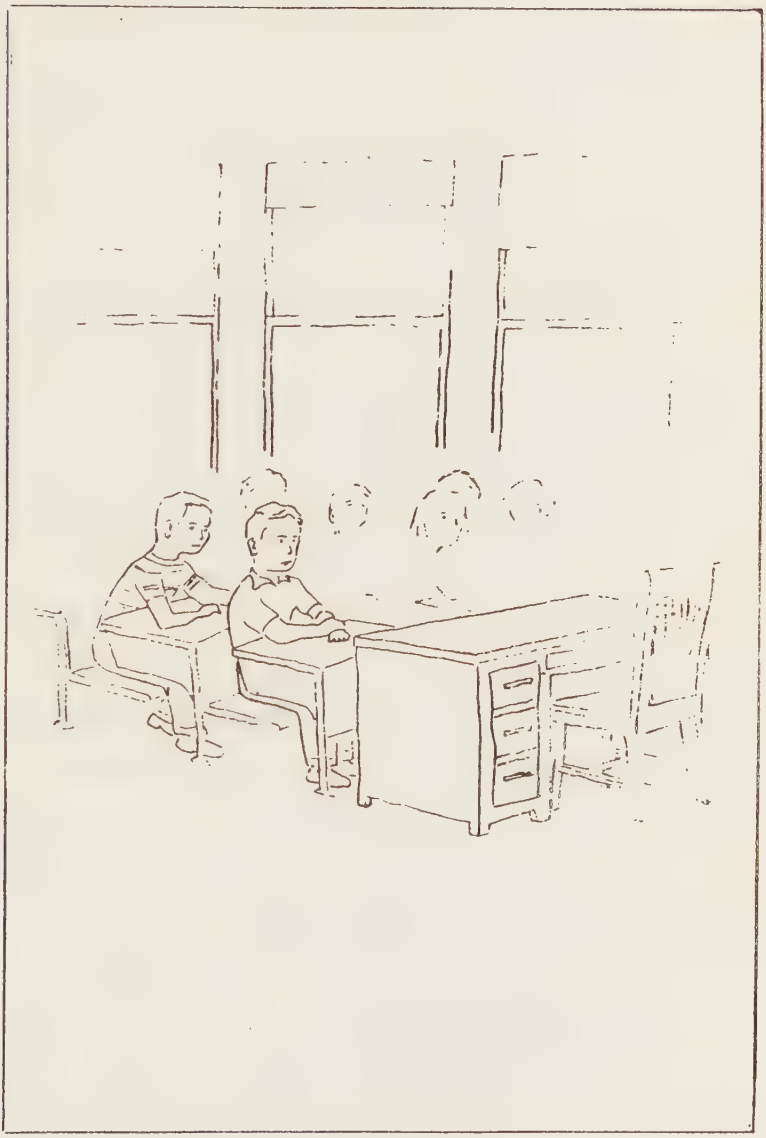












ANHANG II: - Liste der konstruktthematischen und neutralen Aussagen -

WU/HU	
01/ -	Er sagt ihm ordentlich Bescheid; der andere gibt Ruhe.
02/01	Er denkt: "Gleich setzt's was!"
03/ -	Er ist nicht gut aufgelegt.
04/02	Er möchte mit ihm streiten.
05/03	Er ist böse auf ihn.
06/ -	Er gibt ihm eine und denkt: "Die hat er gespürt!"
07/04	Er möchte ihm mal die Meinung sagen.
08/ -	Er spielt ihm übel mit; der andere muß es ausbaden.
09/05	Er denkt: "Dem werd' ich einen Denkzettel verpassen!"
10/ -	Sie geraten aneinander; der andere zieht den kürzeren.
11/ -	Es kommt zu einem ernsthaften Wortwechsel.
12/06	Er denkt: "Dem werd' ich's geben!"
13/07	Er möchte ihn verprügeln.
14/ -	Er denkt: "Der bekommt was zu hören!"
15/ -	Es kommt zu einer Rauferei; der andere unterliegt.
16/ -	Er möchte ihm eine geben, an die er denkt.
17/08	Sie schimpfen miteinander.
18/09	Er sagt ihm die Meinung und denkt: "Geschieht ihm recht!"
19/10	Es kommt zu einem Streit.
20/ -	Er denkt: "Gleich rutscht mir die Hand aus!"
21/ -	Er kann ihn nicht leiden.
22/11	Sie verhauen sich.
23/ -	Er denkt: "Das wird dir noch leid tun!"

24/12 Sie geraten aneinander.

25/ - Er möchte mal in aller Deutlichkeit mit ihm reden.

26/13 Er denkt: "Dem werd' ich gleich eine handfeste Abreibung verpassen!"

27/ - Sie streiten sich; der andere gibt klein bei.

28/ - Er denkt: "Dem werd' ich die Meinung sagen!"

29/14 Es kommt zu einer Auseinandersetzung.

30/ - Er denkt: "Der kann gleich was erleben!"

31/ - Man denkt an die unerfreulichen Folgen einer Auseinandersetzung.

32/15 Es tut ihm leid, daß er nicht gut auf ihn zu sprechen war.

33/16 Er möchte sich wegen des unerfreulichen Vorfalls entschuldigen.

34/ - Er denkt: "Hoffentlich kommt es zu keinem Streit, sonst müßte man sich schämen!"

35/ - Er möchte es nicht im Ernst mit ihm zu tun bekommen.

36/17 Er denkt: "Der meint es garnicht so!", und man einigt sich in Güte.

37/18 Um eine Auseinandersetzung zu verhindern, entschuldigt er sich.

38/ - Nach der Auseinandersetzung quält sie das schlechte Gewissen.

39/19 Sie haben ein schlechtes Gewissen und hören auf, zu streiten.

40/20 Er denkt: "Der meint es garnicht so!", und man einigt sich ohne Streit.

41/ - Er hat Angst vor einem Streit.

42/ - Ein ernsthafter Zwischenfall wird jedoch von außen verhindert.

43/21 Er bereut den Streit.

44/ - Eine Auseinandersetzung wird man hinterher sicher bereuen!

45/ - Er bereut es und versucht, es wieder gutzumachen.

46/22 Er denkt: "Der meint es garnicht so!", und man einigt sich in Frieden.

47/23 Er hat Angst davor, sich mit ihm anzulegen.

48/ - Sie haben ein schlechtes Gewissen und versöhnen sich wieder.

49/24 Er möchte nicht daran schuld sein, daß es zu einem Streit kommt.

50/25 Er möchte die Folgen eines Streits vermeiden.

51/ - Sie werden für die Streitereien bestraft.

52/ - Nach dem Streit quält sie das schlechte Gewissen.

53/26 Er denkt: "Hoffentlich kommt es zu keinem Streit, denn mit dem ist nicht zu spaßen!"

54/27 Um unliebsame Folgen zu vermeiden, läßt er sich nicht auf einen Streit mit ihm ein.

55/ - Eine Auseinandersetzung wird jedoch von außen verhindert.

56/28 Er möchte sich nicht mit ihm einlassen, um ernsthafte Folgen zu verhindern.

- /29 Er möchte etwas mit ihm unternehmen.

- /30 Er macht ihm einen Vorschlag.

- /31 Sie sind gute Freunde.

- /32 Er kann ihn gut leiden.

- /33 Er hilft ihm.

- /34 Er versteht ihn gut.

- /35 Sie treffen sich jeden Tag.

VU: in der Voruntersuchung verwandte Aussagen mit der dort angegebenen Numerierung;

HU: in der Hauptuntersuchung verwandte Aussagen mit der dort angegebenen Numerierung;

AGGRESSIONSMOTIV-Statements: Nr. 01-30 (VU) und Nr. 01-14 (HU)

AGGRESSIONS-HEMMUNGS- " : Nr. 31-56 (VU) und Nr. 15-28 (HU)

PUFFER- " : Nr. 29-35 (HU)

ANHANG III: - ANMERKUNGEN -

(1) Bei der Autoritätsperson handelt es sich um eine Lehrerin, die in einer Schulsituation mit dargestellt ist. Denkbar wäre, daß sie einen spezifischen Hinweisreiz darstellt für

- Aggression gegen Autoritätspersonen und ggfls. gegen deren situative "Korrelate" (Schule, Einrichtungsgegenstände etc.) im Sinne eines generalisierten Aggressionsziels;
- autoritätsbezogene extrinsische Aggressions-Hemmung (insbesondere Strafe durch Erwachsene).

(2) RÖSLER et al. (1980) stellten für die - allerdings geschlechtsindifferent gehaltenen - Personen in den Bildsituationen des LM-GITTERs (SCHMALT 1976 b) der ER-Form die singulare SIE-Form gegenüber und fanden, daß erstere in beiden Populationen (Jungen, Mädchen) durchgehend zu günstigeren psychometrischen Kennwerten führt.

(3) Wir benutzen hier den Begriff "Ambiguität" als Maß(zahl) möglicher, vernünftiger und sich dabei thematisch unterscheidender Apperzeptionsmuster, die im Zusammenhang mit einem Bild gefunden werden (vgl. KENNY & BIJOU 1953; LAZARUS 1953).

(4) Der Autor dankt an dieser Stelle für die konstruktive Mithilfe von Frau Dr. ZUMKLEY-MÜNKEL und Herrn Dr. ZUMKLEY.

(5) Über einen reduzierten Satz von Bildsituationen (im einzelnen mit den Nummern 1, 6, 8, 9 und 11) variieren die Trennschärfen für Aggressions-Statements zwischen .31 und .59, die Trennschärfen für Hemmungs-Statements zwischen .46 und .63. Dies entspricht der erwarteten Trennschärfensteigerung bei reduziertem Anregungsgehalt der Bildsituationen für Aggressionsthematik.

(6) Wir sind hier - in der Entwicklung des alternativen Modells - ein wenig dem Ablauf der Erkenntnisschritte gefolgt, wie sie sich bei uns eingestellt haben. Natürlich liegt dieses Modell schon der unrotierter Zweifaktorenlösung zugrunde (Faktor 1: AGGRESSION; Faktor 2: SOZIALE WERTUNG).

(7) Soweit vorhanden, dürfte sich die Nichteignung des Verfahrens für Mädchen weitgehend auf seinen bildsituativen Teil beziehen, und dort weniger auf den Geschlechtstyp der Handlungsträger (Jungen), als vielmehr auf die mangelnde Repräsentanz der in den Bildern dargestellten Betätigungsfelder im Bereich sozialer Interaktionen für Mädchen.

(8) Wir sprechen insofern von einem "unspezifischen" Faktor Aggressions-Hemmung, als auch die extrinsischen Anteile der Aggressions-Hemmung ihn bedeutsam laden. Die Markiervariablen stammen jedoch aus dem Bereich der intrinsischen Aggressions-Hemmung.

(9) Wir verzeichnen insgesamt nur 2 "faktorielle Sprünge", die erst im vierfaktoriellen Hauptkomponentenmodell des Aggressionsmotiv-Systems zum Vorschein kommen. Theoriekonform konvertiert die Hauptladung der Statement-Variablen Nr. 13 vom Faktor Handlungs- zum Faktor Bedürfnis-bezogener Kognitionen des Aggressionsmotivs und wechselt der Ladungsschwerpunkt der Statement-Variablen Nr. 22 vom Faktor extrinsischer zum Faktor intrinsischer Hemmung von Aggression. Da die Daten der Hauptuntersuchung aussagekräftiger sind, konnte man diese "positiven" Sprünge gewissermaßen erwarten. Wir sahen deshalb darin keinen Hinderungsgrund, diese beiden Statement-Variablen von der Skalenbildung auszuschließen.

(10) Unter Konstanthaltung der psychometrischen Güteeigenschaften waren die entsprechenden Winkel für die DIFF-Skalen nicht mehr frei wählbar, sodaß wir uns bei diesen zugunsten von Rohwerteverteilung und Meßgenauigkeit und zuungunsten der Konfiguration entschieden haben. Ansonsten hätten wir auch dort bestimmte Zeigerstellungen vorgezogen.

(11) Interessant in diesem Zusammenhang ist die Entwicklung der Stabilitätskoeffizienten für einen Kennwert, dessen Berechnung unser GITTER zwar formal gestattet, zu dessen theoretischer Fundierung wir jedoch nicht die Vorkehrungen getroffen haben, wie bei unseren Motiv-Skalen. Es handelt sich dabei um die GITTER-Variable F ("FREUNDSCHAFT"), die angibt, wie oft die Puffer-Statements über alle Bildsituationen hinweg mit JA beantwortet werden. Wir berechnen für F folgende Retestkorrelationen: .78 (2 Wochen), .63 (4 Wochen) und .34 (8 Wochen). Im Gegensatz zu den Motivskalen verzeichnen wir hier einen kontinuierlichen Abfall der Stabilitätskoeffizienten mit zunehmender Dauer des Zeitintervalls zwischen Erst- und Zweitmessung.

(12) Obgleich nach wie vor unklar ist, auf welcher Ebene der PFT eigentlich mißt (z.B. SCHMIDTCHEN 1975, S. 218), so entsprechen doch die von uns verwandten PFT-Variablen ihrer Definition gemäß transsituational konsistenten Persönlichkeitsmerkmalen, wie es dem trait-Konzept entspricht (vgl. WESTERMANN & HAGER 1983), und auch das beim PFT betriebene Validierungsvorgehen (vgl. RAUCHFLEISCH 1979) läßt den Schluß zu, daß mit diesem Verfahren bestimmte, überdauernde Aspekte der Persönlichkeit und weniger des persönlichen Verhaltens in konkreten Situationen erfaßt werden sollen.

(13) Die in unseren Daten aufgefundene Y-Struktur ist offensichtlich PFT-spezifisch. Zunächst entspricht sie in etwa den bei RAUCHFLEISCH (1979, 1, S. 274) aufgeführten Korrelationskoeffizienten zwischen den Signierungsfaktoren E, I und M. Leider sind die entsprechenden Beziehungen zwischen den Kat.-Variablen dort nicht dargestellt. Sodann haben wir sie auch bei LESSER (1959 a, S. 63) gefunden, der mit dem PFT in einer vergleichbaren Population gemessen hat. Jedoch geht auch aus den Angaben von LESSER nicht klar hervor, ob es sich bei den von ihm verwandten PFT-Konstrukta um Kat.-Variablen oder Signierungsfaktoren handelt.

(14) Es ist vorläufig noch als unentschieden zu betrachten, ob im AM-GITTER mehr die operanten oder mehr die respondenten Aspekte des Antwortverhaltens eines Probanden in den Vordergrund treten. Von der Konstruktion her jedenfalls war beabsichtigt, ein operant arbeitendes Verfahren zu erstellen.

(15) Zur faktoriellen Bereinigung der Variablen RAUFEN könnte man eine Umformulierung wie etwa "Fängt mit anderen ohne Grund / absichtlich Raufereien an" in Erwägung ziehen.

ANHANG IV: - Rohwerteverteilungen -

1. Standard-Skalen:

	A	AH	n-A.	Ges.-A.
MEAN	31.69	36.33	-4.65	68.02
MODE	39.00	41.00	-5.00	75.00
MEDIAN	31.59	37.68	-3.75	71.80
STD ERR	0.81	0.87	1.21	1.18
STD DEV	13.67	14.68	20.39	19.72
VARIANCE	186.96	215.49	415.86	389.05
MINIMUM	2.00	1.00	-70.00	8.00
MAXIMUM	76.00	73.00	68.00	126.00
RANGE	74.00	72.00	138.00	118.00
KURTOSIS	0.13	-0.23	0.81	0.37
SKEWNESS	0.29	-0.15	-0.17	-0.40

2. Differentielle Skalen:

	ADM	ADH	AHI	AHE
MEAN	12.95	10.94	13.61	14.57
MODE	10.00	11.00	18.00	15.00
MEDIAN	12.50	10.92	14.14	14.69
STD ERR	0.37	0.32	0.39	0.39
STD DEV	6.12	5.42	6.60	6.47
VARIANCE	37.49	29.41	43.53	41.83
MINIMUM	0.00	0.00	0.00	0.00
MAXIMUM	30.00	29.00	30.00	30.00
RANGE	30.00	29.00	30.00	30.00
KURTOSIS	-0.11	0.21	-0.46	-0.44
SKEWNESS	0.35	0.36	-0.01	0.04

(Haupttuntersuchung, Jungen, n=282; SI/SPSS, FREQUENCIES)

Instruktion zum Rating

Wir haben hier ein Spiel vorbereitet, mit dem wir herausfinden wollen, wie gut Du andere Kinder beobachten kannst.

Du siehst hier ein Blatt, auf dem oben in einem Satz beschrieben wird, was ein anderer Junge tut (Beispiel 1. Satz).

Unter dem Satz siehst Du 7 Kästchen, in denen steht, wie oft Deiner Meinung nach sich ein Junge aus Deiner Klasse so verhält.

Dazu haben wir die Namen Deiner Klassenkameraden auf kleine Kärtchen geschrieben. Deine Aufgabe ist es jetzt, das jeweilige Namenskärtchen in das Kästchen zu legen, wohin es Deiner Meinung nach gehört. Du kannst die Kärtchen so lange hin und her schieben, bis Du denkst, daß sie alle richtig liegen.
Hier bei dem Blatt 1 kannst Du es mal probieren. Wenn Du noch Fragen hast, beantworte ich sie Dir gerne.

1. Er teilt mit anderen, wenn er etwas hat.

1	2	3	4	5	6	7
sehr selten	selten	ziemlich selten	manchmal	ziemlich oft	oft	sehr oft

2. Tut absichtlich nicht, was der Lehrer sagt.

1	2	3	4	5	6	7
sehr selten	selten	ziemlich selten	manchmal	ziemlich oft	oft	sehr oft

3. Traut sich nicht, sich zu wehren, auch wenn er sich ärgert.

1	2	3	4	5	6	7
sehr selten	selten	ziemlich selten	manchmal	ziemlich oft	oft	sehr oft

4. Er hilft anderen.

1	2	3	4	5	6	7
sehr selten	selten	ziemlich selten	manchmal	ziemlich oft	oft	sehr oft

5. Fängt mit anderen Raufereien an.

1	2	3	4	5	6	7
sehr selten	selten	ziemlich selten	manchmal	ziemlich oft	oft	sehr oft

6. Hat ein schlechtes Gewissen, wenn er frech war.

1	2	3	4	5	6	7
sehr selten	selten	ziemlich selten	manchmal	ziemlich oft	oft	sehr oft

7. Wehrt sich, wenn er angegriffen wird.

>

1	2	3	4	5	6	7
sehr selten	selten	ziemlich selten	manchmal	ziemlich oft	oft	sehr oft